

Parameter estimation in coupled ocean-atmosphere general
circulation model using the Ensemble based Filter Method

By

Yun Liu

A dissertation submitted in partial fulfillment of
the requirements for degree of

Doctor of Philosophy

(Atmospheric and Oceanic Sciences)

at the

UNIVERSITY OF WISCONSIN–MADISON

2014

Date of the final oral examination: 6/16/2014

The dissertation is approved by the following members of the Final Oral Committee:

Zhengyu Liu, Professor, Atmospheric and Oceanic Sciences

Daniel J. Vimont, Associate Professor, Atmospheric and Oceanic Sciences

Matthew H. Hitchman, Professor, Atmospheric and Oceanic Sciences

Michael Morgan, Professor, Atmospheric and Oceanic Sciences

Zhigan (Peter) Qian, Professor, Statistics

Shaoqing Zhang, Scientist, Atmospheric and Oceanic Sciences

Acknowledgements

This dissertation has been a long time in the making, and I could not have written it without help from countless people. First and foremost, I would like to thank my advisor, Dr. Zhengyu Liu, for his guidance, patience and support throughout my graduate career. I am also indebted to the members of my committee, Drs. Dan Vimont, Matt Hitchman, Michel Morgan, Zhiguang Qian, and Shaoqing Zhang, for their interest and input in this research.

My gratitude also goes to my research peers and fellow graduate students for their assistance, discussion, collaboration and long-time friendship. My thanks to Drs. Lixin Wu, Robert Jacob, Yafang Zhong, Feng He, Wei Liu, Guangshan Chen, Na Wen, Shu Wu, Xinyao Rong, Jun Chen, Di Ma, Xiaojie Zhu, William Lewis and also to Feiyu Lu, Jiayu Zhang, Yan Yu, Shifan Gu, Jang Zhu, Shan Li, Huaran Liu, Erin Hokanson, Megan Kirchme-young, Ross Dixon, Erin Thomas.

My gratitude also goes to the people at the Center for Climatic Research (CCR) for their great hospitality. My thanks to Drs Lixin Wu, Steve Vavrus, Michael Notaro, John Young, Yi Wang, for their generous help on both scientific and technical issues. I am also grateful to Pat Behling and Kelvin Braun, whose hard work maintaining the computer systems allowed for a smooth computing experience. I wish to thank Melanie Woodworth and Julie Niesen for assistance on numerous paperwork.

My thanks also go to the professors for instructions at the Department of Atmospheric and Oceanic Sciences (AOS). I would also like to thank the staff in the AOS for their excellent service for graduate students. I gratefully appreciate Dr. R. Gallimore,

Miss Megan Kirchmeier and people in UW-Madison writing center for their help on English editing.

Special thanks go to my friends for their generous assistance and making my life joyful as a whole, and those are David Arfa, Sandra Arfa, Xuanji Wang, Zhenglong Li, Wei Zhou, Yinghui Liu, Hong Zhang, Longtao Wu, Fang Wang, Chianyi Liu, Li Bi, Xin Jin, Wenhua Wu, Jun Li, Jun Huang.

Finally, I would like to express my deepest thanks to my wife Juanfang, my children Benjamin and Michelle, my parents, my brother and sister for their tireless love and support.

Abstract

The uncertainties of model parameters are one of the important sources of model bias in a coupled ocean-atmosphere general circulation model (CGCM). The traditional method of tuning parameters through trial-and-error sensitivity experiments is subjective and very inefficient. An ensemble-based filter (EnKF) potentially provides a novel approach for systematic parameter estimation in a CGCM. In this pilot study, we focus on the study of feasibility and effectiveness of parameter estimation in a CGCM using an EnKF strategy.

We developed a random subgrouping scheme to improve the filter performance of EnKF in a nonlinear regime. EnKF is a linear filter, while model response to a parameter uncertainty could be very nonlinear. Therefore, it is important to improve the ability of an EnKF in handling nonlinearities. The random subgrouping scheme alleviates the outlier effect occurring during EnKF assimilation in a nonlinear system. Test results, using the random subgrouping technique on two low-order models and one intermediate model, show that the new scheme significantly improves performance compared to regular EnKFs.

In the twin experiment framework, we perform the first ensemble based parameter estimation in a CGCM. We first perform single parameter estimation and then multiple-parameter estimation. Results show that the biased parameters all converge to the “truth” values after the assimilation of the oceanic observations of monthly sea surface temperature and salinity and atmospheric data of temperature and wind (only for multiple-parameter estimation). Finally, the improved parameters also improve the model climatology.

An adaptive spatial average (ASA) algorithm is proposed to increase the efficiency of parameter estimation. For a complex system such as a CGCM, the sensitivity and response of a model variable to a model parameter could vary spatially and temporally. Refined from a previous spatial average method, the ASA uses the ensemble spread as the criterion for selecting “good” values from the spatially varying posterior estimated parameter values; the “good” values are then averaged to give the final globally uniform posterior parameter. In comparison with existing methods, the ASA parameter estimation shows superior performance: faster convergence and enhanced signal-to-noise ratio.

Table of Content

Acknowledgement	i
Abstract	iii
Table of Content	v
Chapter 1 Background	1
1.1 Introduction	1
1.2 Data assimilation	3
1.2.1 Kalman Filter (KF)	6
1.2.2 Ensemble Based Filters (EnKF)	7
1.3 Ensemble based parameter estimation	8
1.4 Techniques for Parameter ensemble	10
1.4.1 Parameter ensemble spread inflation	10
1.4.2 The correlation cutoff scheme for parameter estimation	11
1.4.3 “ Spin-up” parameter estimation	13
1.5 Parameter estimation example in a simple model	13
1.6 Thesis outline	15
Chapter 2 A Random Subgrouping Scheme for Ensemble Based Filters ...	17
2.1. Introduction	19

2.2 The Algorithm of Ensemble Based Filters	21
2.2.1 Stochastic filter (EnKF)	21
2.2.2 Deterministic filter (EnSRF)	21
2.2.3 The Random Subgroup Ensemble Based Filters	22
2.3 Results	26
2.3.1 sEAKFn in the Lorenz-63 model	26
2.3.2 sEAKFn in Lorenz-96 model – model independence	35
2.3.3 sEnSRF in QG-slab model	37
2.4 Summary	39
2.5 Appendix 2A: Lorenz-63 model	40
2.6 Appendix 2B : Lorenz-96 model	41
2.7 Appendix 2C: QG-slab model	42
Chapter 3 Ensemble-based Parameter Estimation in a Coupled General Circulation Model	55
3.1 Introduction	57
3.2 Model and Methodology	57
3.2.1 The Fast Ocean Atmosphere Model (FOAM)	57
3.2.2 Coupled data assimilation with enhanceive parameter correction (DAEPC)	58

3.2.3 The Adaptive Spatial Average scheme (ASA)	58
3.2.4 Observations and ensemble configuration	60
3.3 Single parameter estimation	62
3.4 Multiple parameter estimation	65
3.5 Summary	69
3.6 Appendix: An experiment with unsuccessful parameter estimation	71

Chapter 4 Ensemble-based parameter estimation in a coupled GCM using the Adaptive Spatial Average method	82
4.1 Introduction	84
4.2 Model and Method	85
4.2.1 Fast Ocean Atmosphere Model (FOAM)	85
4.2.2 Data Assimilation Scheme	85
4.3 Model sensitivity with respect to solar penetration depth	87
4.4 The Adaptive Spatial Average scheme (ASA)	89
4.5 Comparison of ASA with GPO and SA	91
4.5.1 The assimilations with GPO scheme	91
4.5.2 Comparison between SA and ASA	94
4.6 Summary and Discussions	96
4.7 Appendix: Preliminary Theoretical Consideration for ASA	98

Chapter 5 Summary and Future work	112
5.1 Summary	114
5.2 Concluding remarks	114
5.3 Future work	115
References	117

Chapter 1 Background

1.1 Introduction

In spite of tremendous efforts made by the climate modeling community over decades, current climate model still suffer from significant climate biases. One important source of bias is the uncertainty of model parameters. The tuning of model parameters, however, has remained a challenging task, especially in a complex climate model such as a coupled ocean-atmosphere general circulation model (CGCM), which consists of nonlinear dynamic processes on a wide range of time scales and requires substantial computational resources. For example, model biases occur when the atmosphere and ocean components are coupled, indicating poor representations of atmospheric and oceanic processes associated with coupling. One major bias is the so-called tropical bias, characterized by a too westward penetration of the cold tongue and a double ITCZ in the Pacific and Atlantic sectors (e.g. Mechoso et al., 1995; Davey et al., 2002; Collins et al., 2006; Lin, 2007). Tuning model parameters in a single component model (atmospheric or oceanic) individually cannot alleviate this type of coupled bias. The atmosphere model and ocean model can usually be tuned individually in the stand-alone mode reasonably well. Once they are coupled, however, the coupled model drifts away significantly from the observation.

Traditionally, modelers tune the model parameters in a complex CGCM subjectively through trial-and-error sensitivity experiments. This subjective tuning is inefficient because of the huge computational costs. Furthermore, given the large number of parameters in a CGCM the limitation of a human mind and the intensive labor involved, it is virtually impossible to identify the optimal setting of parameters using the subjective tuning alone. On the other hand, modern data assimilation approaches such as the four-dimension variational assimilation (4D-Var) and Ensemble-based filter (EnKF) provide the potential for an automatic optimization of the model

parameters along with the model states (Navon, 1997; Derber, 1989; Anderson, 2001). In practice, however, it has remained unclear if there is a data assimilation strategy practical for parameter optimization in a complex system like the CGCM.

For a complex system as a CGCM, the parameter estimation with EnKF is much more practical than with 4D-VAR because the EnKF is easy to implement and maintain relative to the 4D-VAR scheme. The EnKF is an off-line scheme and therefore can be adapted to any model configuration conveniently. The 4D-VAR method is an on-line scheme, which requires extraordinary effort to build and update an adjoint model (Navon, 1997; Zou et al., 1992; Zhu and Navon, 1998), which remains a great challenge for any complex CGCM. The implementation of 4D-VAR methods in a complex CGCM still remains a great challenge, even for traditional data assimilation. The first EnKF system has been implemented successfully in a CGCM for climate applications (Zhang et al., 2007, 2008) while, so far, no 4DVAR scheme has been shown to work in a CGCM, although it does work in the operational numerical weather prediction (NWP) models.

Previous studies in ensemble-based parameter estimation have shown encouraging results. There are different approaches to do parameter estimation using EnKF. One approach focuses on the improvement of model parameters in order to improve steady-state climatology by assimilating the long-term climatological observations iteratively. This approach has been applied to the highly non-linear Lorenz (1963) system (Annan and Hargreaves, 2004), as well as a climate model of intermediate complexity (Annan et al., 2005a), an atmosphere NWP model (Annan et al., 2005b), and an earth system model of intermediate complexity (Ridgwall et al., 2007). Another approach focuses on the improvement of model parameters as well as model forecast by implementing a simultaneous estimation of model state variables and parameter with

the assimilation of time evolutionary observations. This approach has been applied to mesoscale atmospheric models, including a planetary boundary layer model (Hacker and Snyder, 2005), a 2-D sea breeze model (Aksoy et al., 2006a), a regional atmosphere NWP model (MM5, Aksoy et al., 2006b), an advanced regional weather prediction system (Tong and Xue, 2008 a & b), and a regional atmosphere NWP model using realistic observation (WRF, Hu et al., 2010). The approach of simultaneous estimation of model state and parameter has also implemented in the coupled climate models, including conceptual coupled climate models (Zhang, 2011 a & b; Zhang et al., 2012), and an intermediate coupled atmosphere-ocean-land model (Wu et al. 2012, a & b).

In this study, we will explore the ensemble based parameter estimation in a CGCM using the simultaneous parameter and state estimation approach. As a pilot study, we will focus on the study of feasibility and effectiveness of parameter optimization. This study will give us experience in parameter estimation in a CGCM and will lay a foundation for future studies towards a systematic improvement of complex coupled earth system models. It will have a significant impact on the development of complex coupled models in the future.

1.2 Data assimilation

Data assimilation is the technique that incorporates observational data into a compute model of a real system. It proceeds by analysis cycles that consist of steps: an analysis step and a forecast step. In each analysis step, the observations and the model forecast (also called the prior or background) of the system are combined together to produce the analysis (all called the posterior), which is considered as 'the best' estimate of the current state of the system. In each forecast step, the model is then advanced in time with the analysis as the initial condition and its result becomes the forecast in the next analysis cycle. An analysis itself can be as a

comprehensive and self-consistent diagnostic product. It can also use as the initial condition for a model forecast.

Due to uncertainties in climate modeling, the temporal evolution of model climate states can be described as a continuous stochastic dynamical process through a set of stochastic differential equations (Jazwinski 1970)

$$\frac{dx_t}{dt} = \mathbf{F}(\mathbf{x}_t, t) + \mathbf{G}(\mathbf{x}_t, t)\mathbf{W}_t \quad (1.1)$$

Here, \mathbf{x}_t is the vector for the model state variable at time t , and \mathbf{F} is a vector function representing the deterministic part of the model, including its dynamical and physical processes. The last term at the right hand side of eqn. (1.1) represents the contributions of uncertainties resulted from erroneous initial conditions with \mathbf{W}_t as a white (in time) Gaussian process and \mathbf{G} as a matrix that defines the relationship between \mathbf{W} and \mathbf{x}_t . An observation vector \mathbf{y}_t at time t can be written as

$$\mathbf{y}_t = \mathbf{h}(\mathbf{x}_t, t) + \mathbf{v}_t \quad (1.2)$$

where \mathbf{h} is the mapping function from the model state space (\mathbf{x}_t) to the observed space (\mathbf{y}_t); \mathbf{v}_t is the observation error. Generally, the model states \mathbf{x}_t and observations \mathbf{y}_t described by probability distribution density functions (PDF). Data assimilation solves the conditional PDF of model states at time t given all the prior observations. Here the conditional PDF indicates as $p(\mathbf{x}_t|\mathbf{Y}_t)$, where $\mathbf{Y}_t = [\dots, y_{t-1}, y_t]$ is the set of all observations that are taken up to and including time t . Following the derivation in (Jazwinski 1970, Anderson and Anderson 1999, Anderson 2001), based on Bayes' rule, the $p(\mathbf{x}_t|\mathbf{Y}_t)$ can be derived as

$$p(\mathbf{x}_t|\mathbf{Y}_t) = p(\mathbf{x}_t|\mathbf{y}_t, \mathbf{Y}_{t-1}) = \frac{p(\mathbf{y}_t|\mathbf{x}_t, \mathbf{Y}_{t-1})p(\mathbf{x}_t|\mathbf{Y}_{t-1})}{p(\mathbf{y}_t|\mathbf{Y}_{t-1})} \quad (1.3)$$

$$p(\mathbf{x}_t|\mathbf{Y}_{t-1}) = \frac{p(\mathbf{x}_t, \mathbf{Y}_{t-1})}{p(\mathbf{Y}_{t-1})} \quad (1.4)$$

Incorporating (1.4) into (1.3) gives

$$p(\mathbf{x}_t|\mathbf{Y}_t) = \frac{p(\mathbf{y}_t|\mathbf{x}_t, \mathbf{Y}_{t-1})p(\mathbf{x}_t|\mathbf{Y}_{t-1})}{p(\mathbf{y}_t|\mathbf{Y}_{t-1})} \quad (1.5)$$

The $p(\mathbf{y}_t|\mathbf{x}_t, \mathbf{Y}_{t-1})$ is the PDF of observation \mathbf{y}_t and the $p(\mathbf{x}_t|\mathbf{Y}_{t-1})$ is the PDF of prior. The denominator $p(\mathbf{y}_t|\mathbf{Y}_{t-1})$ is a normalization that guarantees that the total probability of all possible states is 1.

For linear data assimilation problem, eqn. (1.5) can be simplified with the assumption of Gaussian PDFs for both prior and observations. A Gaussian PDF is completely determined by the mean and covariance,

$$p(\mathbf{x}_t|\mathbf{Y}_{t-1}) \sim N(\mathbf{x}^f, \mathbf{P}^f) \propto e^{-\frac{1}{2}(\mathbf{x}_t - \mathbf{x}^f)^T \mathbf{P}^{f-1}(\mathbf{x}_t - \mathbf{x}^f)} \quad (1.6)$$

$$p(\mathbf{y}_t|\mathbf{x}_t, \mathbf{Y}_{t-1}) \sim N(\mathbf{y}_t, \mathbf{R}) \propto e^{-\frac{1}{2}(\mathbf{y}_t - \mathbf{h}(\mathbf{x}_t))^T \mathbf{R}^{-1}(\mathbf{y}_t - \mathbf{h}(\mathbf{x}_t))} \quad (1.7)$$

Here \mathbf{x}^f and \mathbf{y}_t are the prior and observations and \mathbf{P}^f and \mathbf{R} denote their error covariance matrixes, respectively; N denotes the normal distribution. Incorporating (1.6) and (1.7) into (1.5), the $p(\mathbf{x}_t|\mathbf{Y}_t)$ also follows Gaussian distribution with

$$p(\mathbf{x}_t|\mathbf{Y}_t) \propto e^{-\frac{1}{2}[(\mathbf{y}_t - \mathbf{h}(\mathbf{x}_t))^T \mathbf{R}^{-1}(\mathbf{y}_t - \mathbf{h}(\mathbf{x}_t)) + (\mathbf{x}_t - \mathbf{x}^f)^T \mathbf{P}^{f-1}(\mathbf{x}_t - \mathbf{x}^f)]} \quad (1.8)$$

From (1.8), we can get the basic cost function for linear data assimilation

$$J(\mathbf{x}_t) = \frac{1}{2} [(\mathbf{y}_t - \mathbf{h}(\mathbf{x}_t))^T \mathbf{R}^{-1}(\mathbf{y}_t - \mathbf{H}\mathbf{x}_t) + (\mathbf{x}_t - \mathbf{x}^f)^T \mathbf{P}^{f-1}(\mathbf{x}_t - \mathbf{x}^f)] \quad (1.9)$$

where \mathbf{H} is the linearized mapping function (forward operator) from state variables space to observation space (see eqn. (1.2)). The $J(\mathbf{x}_t)$ is a quadratic form of \mathbf{x}_t and therefore the $p(\mathbf{x}_t|\mathbf{Y}_t)$ follows a Gaussian distribution.

$$p(\mathbf{x}_t|\mathbf{Y}_t) \sim N(\mathbf{x}^a, \mathbf{P}^a) \propto e^{-\frac{1}{2}(\mathbf{x}_t - \mathbf{x}^a)^T \mathbf{P}^{a-1}(\mathbf{x}_t - \mathbf{x}^a)} \quad (1.10)$$

Therefore we can solve the conditional distribution of $p(\mathbf{x}_t|\mathbf{Y}_t)$ by find the mean \mathbf{x}^a are and covariance \mathbf{P}^a of analysis \mathbf{x}_t .

The three-dimension variational assimilation scheme (3D-Var), Kalman filter (KF, Kalman, 1960; Kalman and Bucy, 1961) and EnKF can all start from the cost function of (1.9). The differences among these methods are the way they deal with the background covariance matrix P^f . The 3D-Var assumes the P^f to be stationary, while KF and EnKF use a flow-dependent P^f .

1.2.1 Kalman Filter (KF)

The KF (Kalman 1960; Kalman and Bucy 1961) achieves an optimal estimation of x^a, P^a in (1.10) for either the condition of minimum variance or the condition of maximum likelihood based on the assumption of Gaussian (or normal) distributions for both forecast and observation.

The KF updates the analysis (x^a) and covariance (P^a) as

$$x^a = x^f + K(y^o - Hx^f) \quad (1.11)$$

$$P^a = (1 - KH)P^f \quad (1.12)$$

where $K = P^f H^T (HP^f H^T + R)^{-1}$ is the Kalman gain. H is the linearized mapping function from state variables space to observation space, $y = Hx$ and H^T is the transpose of H. The observational uncertainty R is determined by the error scale of observations. The forecast uncertainty P^f is advanced from the previous P^a by using the Fokker–Planck Kolmogorov equation (Jazwinski 1970),

$$\frac{dP_t}{dt} = FP_t + P_t F^T + GQG \quad (1.13)$$

where P_t indicates the uncertainty of P^f or P^a (separated by time t). The variable Q is covariance of W_t . When x_t has the dimension of M , the P_t has the dimension of M^2 . Therefore it is difficult to apply (1.13) in high dimensional systems like weather or climate models.

1.2.2 Ensemble Based Filters (EnKF)

The EnKF is a Monte Carlo implementation of KF (Evensen, 1994; Burger et al., 1998; Houtekamer and Mitchell, 1998; Houtekamer et al., 2005; Anderson, 2001, 2003, Bishop et al. 2001, Whitaker and Hamill 2002, Tippett et al 2003, Sakov and Oke 2007). The original KF need to advance the covariance matrix P in time. However, it is not feasible computationally to maintain and update the covariance matrix in high-dimensional systems. For practical solution, researchers have developed EnKFs. An EnKF represents the PDF of the model state using a collection of state vectors, called an ensemble, and replace the background and analysis covariance matrix P^f and P^a by the sample covariance P_e^f and P_e^a , respectively. The model state variables are updated by observations through EnKF as

$$\mathbf{x}^a = \mathbf{x}^f + \mathbf{K}_e(\mathbf{y}^o - \mathbf{H}\mathbf{x}^f) \quad (1.14)$$

$$\mathbf{P}_e^a = (\mathbf{1} - \mathbf{K}_e\mathbf{H})\mathbf{P}_e^f(\mathbf{1} - \mathbf{K}_e\mathbf{H})^T + \mathbf{K}_e\mathbf{R}\mathbf{K}_e^T = (\mathbf{1} - \mathbf{K}_e\mathbf{H})\mathbf{P}_e^f \quad (1.15)$$

Now the Kalman gain $\mathbf{K}_e = \mathbf{P}_e^f\mathbf{H}^T(\mathbf{H}\mathbf{P}_e^f\mathbf{H}^T + \mathbf{R})^{-1}$ is a suboptimal solution for \mathbf{K} because it is estimated from a finite size ensemble.

EnKF uses ensembles to sample model uncertainties of forecast or analysis. The ensemble spreads quantify the uncertainties. A distorted ensemble spread could decrease the quality of analysis significantly. Especially when a forecast ensemble spread is much smaller than the true uncertainty of the forecast, the EnKF analysis scheme tends to overweight the forecast and ignore the impact of observation, and therefore causes filter divergence. Based on the methods of updating analysis ensemble, EnKF can be divided into two categories: stochastic and deterministic. We will discussion difference between them in the chapter 2.

1.3 Ensemble based parameter estimation

Parameter estimation deals with a very different problem from that of the state variable estimation, although it uses the same updating equation as state variables. Parameter estimation is a smoothing problem while state estimation with EnKF is a filtering problem. Given a realization of the observation vector of $\mathbf{Y}_t = [\dots, \mathbf{y}_{t-1}, \mathbf{y}_t]$, an estimation problem compute an estimation of \mathbf{x}_k based on \mathbf{Y}_t . If $k < t$, the problem is called a smoothing problem; if $k = t$, it is called filtering problem; if $k > t$, it is a prediction problem. Generally we evaluate an estimated parameter related to its initial guess value. In that sense parameter estimation is a smoothing problem, while the state estimation with EnKF is a filter problem, which minimizes the uncertainty of model state at current time.

In data assimilation, parameters can be estimated by augmenting state variables with model parameters (e.g. Banks, 1992a & b; Anderson, 2001). The parameters are treated as special variables that are stationary during model integration. When we also consider the uncertainty of parameters in climate modeling, eqn. (1.1) and (1.5) are updated respectively, as

$$\frac{d\mathbf{x}_t}{dt} = \mathbf{f}(\mathbf{x}_t, \boldsymbol{\beta}_t, t) + \mathbf{G}(\mathbf{x}_t, \boldsymbol{\beta}_t, t)\mathbf{W}_t \quad (1.16)$$

$$\frac{d\boldsymbol{\beta}_t}{dt} = 0 \quad (1.17)$$

$$p(\mathbf{x}_t, \boldsymbol{\beta}_t | \mathbf{Y}_t) = \frac{p(\mathbf{y}_t | \mathbf{x}_t, \boldsymbol{\beta}_{t-1}, \mathbf{Y}_{t-1})p(\mathbf{x}_t, \boldsymbol{\beta}_{t-1} | \mathbf{Y}_{t-1})}{p(\mathbf{y}_t | \mathbf{Y}_{t-1})} \quad (1.18)$$

Here, $\boldsymbol{\beta}_t$ is the vector for the parameters at time t. Here $\boldsymbol{\beta}_{t-1}$ represents the estimated parameter vector given all previous observations (\mathbf{Y}_{t-1}) and $\boldsymbol{\beta}_t$ represents the adjusted parameter vector when the new observation (\mathbf{y}_t) is assimilated. Now $p(\mathbf{y}_t | \mathbf{x}_t, \boldsymbol{\beta}_{t-1}, \mathbf{Y}_{t-1})$ is the PDF of observation \mathbf{y}_t and $p(\mathbf{x}_t, \boldsymbol{\beta}_{t-1} | \mathbf{Y}_{t-1})$ is the PDF of prior. Therefore, based on the assumption of Gaussian (or normal) distributions for both state variables (\mathbf{x}_t) and parameters ($\boldsymbol{\beta}_t$), EnKF can

apply the parameter estimation use the same equations as the traditional EnKF in (eqn. (1.14) and (1.15)), except using the augmented vectors of $\begin{bmatrix} x^f \\ \beta^f \end{bmatrix}$ and $\begin{bmatrix} x^a \\ \beta^a \end{bmatrix}$ as the prior and posterior, respectively. Here the β^a and β^f are the ensemble of prior and updated parameters, respectively. In this manner, the model states and parameters can be estimated simultaneously. The state variables remain updated the same as in the traditional EnKF assimilation while the parameters are updated as

$$\beta^a = \beta^f + C(\beta^f, Hx^f)(HP^f H^T + R)^{-1}(y - Hx^f) \quad (1.19)$$

The key to parameter correction is the covariance between the parameters and the simulated observation $C(\beta^f, Hx^f)$, which reflects the sensitivity of model simulated observation to model parameters. When the parameter-state covariance is signal-dominated and the observation sufficiently constrains the model states, it is possible to correct these parameters with EnKF.

The ensemble-based filter updates the parameter using the same manner as the traditional EnKF updating state. Therefore, successful implementation of EnKF for pure data assimilation purpose is the precondition for ensemble based parameter estimation. A sufficiently constrained simulated observation enhances the signal/noise ratio of parameter-state covariance. The EnKF is a linear filter, although it could be used in the nonlinear regime. While a model response to a parameter uncertainty could be very nonlinear, we need to consider the effect of nonlinearity, especially for parameter estimation purpose. The stochastic and deterministic EnKFs perform very differently in a nonlinear system. In chapter 2, we will investigate the differences in simple nonlinear systems and propose a random subgrouping scheme for deterministic EnKFs to improve the assimilation performance under nonlinear conditions.

1.4 Techniques for Parameter ensemble

In general, however, there is no guarantee that the optimal parameters can be identified from state variables using data assimilation (Navon, 1998) especially for a complex nonlinear system such as a CGCM. Parameter estimation could fail if model responses to parameter uncertainty is not sufficiently strong on the observation space to overcome the sampling errors or the model response to parameter uncertainty does not vary smoothly (Navon, 1997; Zupanski and Zupanski, 2006; Nielsen-Gammon, 2010). Here we will address some techniques to increase the probability of success of parameter estimation.

1.4.1 Parameter ensemble spread inflation

Parameter estimation is an inverse problem. The inverse problem is considered the "inverse" to the forward problem, that is, the errors of model forecast or simulation related the model parameters error. A biased parameter could generate model state bias on the mean state and/or its variation. The state bias, reflected in the model forecast, is the signal for the parameter estimation. EnKF maps the signal from observation space onto parameter space to correct the parameter error. To obtain the mapping function, EnKF uses a parameter ensemble to sample the parameter bias. The parameter ensemble is then applied in the model integration to obtain the response of model forecasts to parameter ensemble. The covariance between the forecast ensemble and parameter ensemble samples the mapping function.

One concern for parameter estimation is the spread (standard derivation) of parameter ensemble at initialization and during the data assimilation process. Parameter ensemble plays critical role in generating parameter-state covariance for parameter estimation as well as enhances the forecast state ensemble and improves the state estimation because the forecast ensemble includes the uncertainty generated by parameter error. However, model parameter

initial errors are unknown in the real world. No straightforward guidance exists for the proper range of ensemble spreads of parameters to be estimated. Furthermore, parameters are not dynamical variables. The model integration at forecast steps can only increase the ensemble spreads of model states but not parameters ensemble spread. The parameter ensembles remain constant during forecast steps. While the variance of both model variables and parameters is reduced at an analysis step. As a result, parameter spread remains unchanged between analysis steps only to be reduced again during the following analysis step. This leads to a progressively decreasing parameter ensemble spread. The progressively decreasing parameter eventually could cause too small of a parameter ensemble spread and filter divergence for parameter estimation, especially if the mean parameter value does not promptly converge toward the true value or if the true parameter value varies in time.

In a twin model experiment, the initial ensemble spread for a parameter can be set as the initial error of the parameter (absolute difference between initial mean and true parameter value). Usually, parameter ensemble spread is inflated during assimilation to prevent filter divergence. For example, the conditional covariance inflation technique inflates (CCI) the parameter ensemble back to a predefined minimum value when necessary (Aksoy et al. 2006b); Hu et al (2010) inflates the parameter ensemble back to the initial ensemble spread. The CCI is used in our parameter estimation experiments below.

1.4.2 the correlation cutoff scheme for parameter estimation

Ensemble based parameter estimation uses parameter-state covariance to represent the mapping function from observation space to parameter space. Successful parameter estimation requires the covariance to be signal-dominant. When we use an observation y^o and a forecast x^f

to update a parameter β , the β is augmented with the state variable x . Thus, the updating equations are the same as the traditional EnKF equations. The uncertainty scale of simulated observation x^f can be written as

$$\sigma_x^2 = L^2 \sigma_\beta^2 + \xi_\beta^2 + \sigma_{ib}^2 \quad (1.20)$$

where σ_x, σ_β are the error scales (standard derivations) of x^f and β respectively.

The first two terms at the right hand side of equation are the uncertainty produced by the parameter uncertainty. The uncertainty has been separated into linear partition ($L^2 \sigma_\beta^2$) and nonlinear partition (ξ_β^2). L is the flow dependent linear response coefficient of x to parameter β . The third term (σ_{ib}^2) is the partition of uncertainty related to other sources, like the initial condition uncertainty, boundary condition uncertainty, and model bias, which are not related to the parameter uncertainty. Only $L^2 \sigma_\beta^2$ is the signal for parameter estimation in a linear filter system.

One major noise source for parameter estimation with EnKF is the sampling error generated from limited ensemble size. The noise scale of forecast is proportional to the scale of simulated observation uncertainty and decreases when the degrees of freedom (df) increase. Here it is written as

$$\sigma_n^2 = \kappa(df) \sigma_x^2 \quad (1.21)$$

where $\kappa = \kappa(df)$ is a positive coefficient related to the df .

Generally, we can enhance the signal/noise ratio, and therefore improve the parameter estimation from three aspects: (i) increase the ensemble size to decrease the sampling error, which is constrained by the computational resource; (ii) decrease the uncertainties of forecast by produce a high quality analysis, which is constrained by the observation quality and coverage; (iii) increasing the weight of the parameter ($L^2 \sigma_\beta^2$) among the total forecast uncertainty σ_x^2 in

equation (1.20). Usually the signal for parameter estimation and model forecast uncertainty vary temporally. We can further enhance the signal-to-noise ratio by only applying parameter updating under the condition with the signal weighted among model forecast uncertainty. That can be easily achieved with a correlation cutoff scheme. The correlation cutoff scheme only allows the observation update parameter when the value of parameter-state correlation is bigger a predefined cutoff value. The value of parameter-state correlation indicates the weight of parameter signal among total forecast uncertainty.

1.4.3 “Spin-up” parameter estimation

To enhance the signal/noise ratio of parameter-state correlation, Zhang et al. (2012) proposed a “spin-up” process for parameter estimation. Before the parameter estimation is activated, a “spin-up” process for the state estimation to reach a “quasi-equilibrium” state such that the uncertainty of model states is sufficiently constrained by observations. The “spin-up” process decreases the forecast uncertainties related to the initial condition (the σ_{ib}^2 in eqn. (1.20)). This process also provides the sufficient time for the model to respond the parameter ensembles.

1.5 Parameter estimation example in a simple model

In this section, we will offer an example of parameter estimation in a conceptual coupled “climate” model – Lorenz63-slab model: a slowly-varying variable T coupled with the 3-variable Lorenz model (Lorenz 1963) forms a conceptual “atmosphere” coupling with 1-variable slowly varying slab-ocean

$$\frac{dx}{dt} = -\sigma(x - y) \quad (1.22)$$

$$\frac{dy}{dt} = -xz + (1 + c_1 T)\beta x - y \quad (1.23)$$

$$\frac{dz}{dt} = xy - cz \quad (1.24)$$

$$h \frac{dT}{dt} = c_2 y - \lambda T + S(t) \quad (1.25)$$

$$c_1 = \begin{cases} 0 & y \leq c_0 \\ 0.05 & y > c_0 \end{cases} \quad (1.26)$$

where x, y and z are the high-frequency “atmosphere” variables with the original σ, β and c parameters. The “atmosphere” sustains the chaotic nature with the standard values of σ, β and c (10, 28, $8/3$, respectively). The “atmosphere” goes through a lobe of the Lorenz attractor in approximately 1 non-dimensional time unit, around 10 days in the real world). The T represents a slowly varying “ocean” temperature; the big ocean thermal capacity is indicated by $h = 10$. The simplest slab ocean consists of a linear damping term λT and an imposed external forcing $S = 10 + \cos\left(\frac{2\pi t}{10}\right)$. The c_1 and c_2 are the coupling coefficients between the “atmosphere” and “ocean”. The model is integrated using a 4-th order Runge-Kutta method with a time resolution of $dt = 0.01$. A similar “toy” model is applied by Zhang et al (2012). The major difference is that we add a threshold-type parameter c_0 to modulate c_1 (eqn. (1.25)).

Here we choose σ, β, c, c_2 as the parameters for estimation. The initial values of these estimated parameters are 20% larger than their true values. The ensemble size is 30 and the observations are daily (x, y, z, T) with error scales of (2, 2, 2, 0.2). A multiple parameter estimation with EnKF combined with the techniques in section 1.5 leads to a successful estimation of the parameters (Fig.1.1). All four parameters converge to the “truth” values after two years of assimilations. The parameters of c_2 converges slower than the other three parameters because it is the parameter related to the slow ocean component, while the other three parameters of σ, β, c are related to the fast atmospheric component.

1.6 Thesis outline

This work is devoted to develop an effective coupled data assimilation strategy for parameter optimization in a CGCM using EnKF strategy (Evenson, 1994, 2003). Due to the lack of previous work and experience on parameter estimation in CGCMs, We will investigate the feasibility of parameter estimation in a CGCM under a twin model framework, where the biased parameters are the only model error sources.

The thesis is organized as follows. We will propose a random subgrouping scheme of EnKF for deterministic EnKFs in chapter 2. The random subgrouping technique can significant improve the performance of a deterministic EnKF under nonlinear conditions. We will demonstrate the first successful parameter estimation a CGCM using a refined technique - the adaptive spatial average (ASA) algorithm in chapter 3. A detail discussion of the ASA algorithm is given in chapter 4. A summery and description of future works are given in chapter 5.

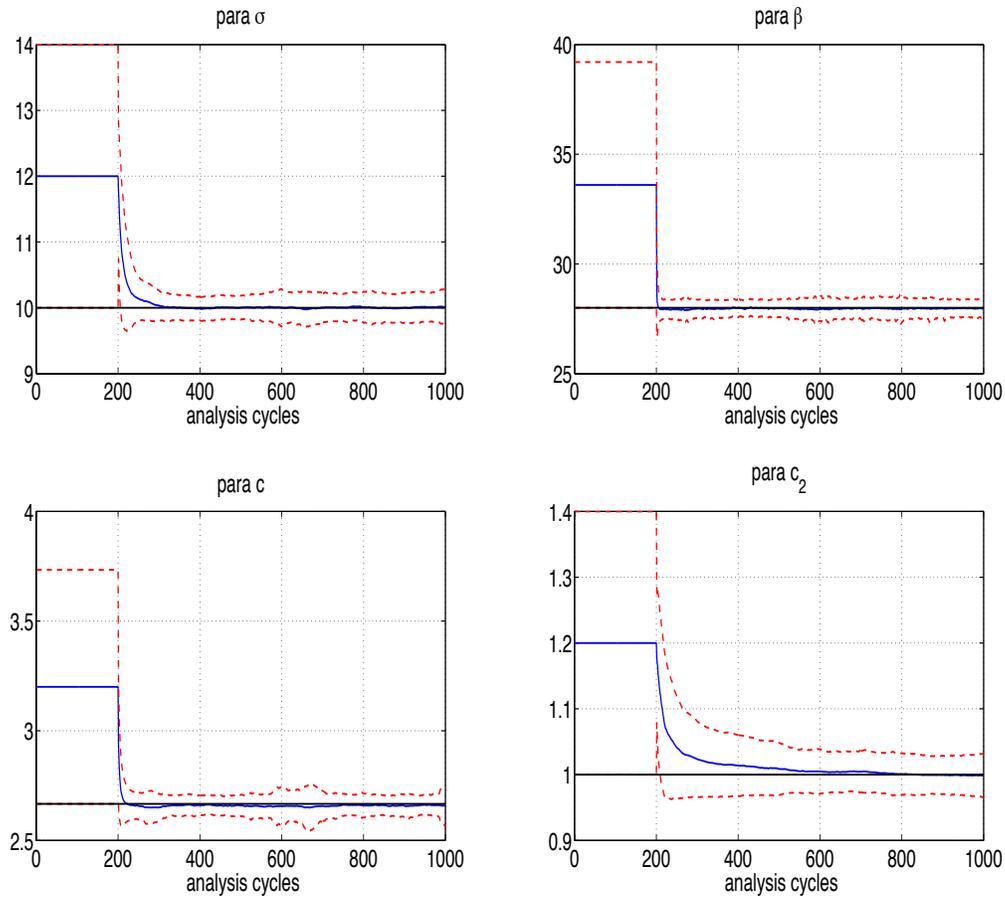


Figure 1.1 The multiple parameter estimation of σ , β , c , c_2 in Lorenz63-slab model. The ensemble size is 30. The observations are (x, y, z, T) with the error scale of $(2, 2, 2, 0.2)$, respectively. The blue lines are for the temporal evolution of the ensemble mean of parameters and the red dashed lines are the 1-standard deviation of the ensemble spread. The black solid lines are the “truth” and the black dashed lines are the minimum parameter ensemble spreads (uncertainty goals) of CCI in the experiment. The first 200 analysis cycles are the “spin-up” period before the parameter estimation is activated.

Chapter 2: A Random Subgrouping Scheme for Ensemble Based Filters

The ensemble based parameter estimation is based on the data assimilation skill of EnKF. Therefore a high quality of an EnKF implementation is the precondition for ensemble based parameter estimation. An EnKF is optimal only for linear systems because it is a linear filter. Filter divergence could take place when it is used in the nonlinear regime (Evensen 1997; Lawson and Hansen 2004). While the model response to a parameter uncertainty could be very nonlinear. Therefore, it is important to consider the ability of an EnKF in handling nonlinearities. The higher performance of an EnKF in a nonlinear system provides the higher quality for parameter estimation.

Abstract

Ensemble based filters can be divided into two categories: stochastic and deterministic. Both types of filters suffer from the problem of generating extreme outliers in the ensembles produced in a nonlinear system. This is especially true for the deterministic filter with a large ensemble size. The outliers can persist for a long time, generate substantial separation from the other ensemble members, and contribute to large errors in the ensemble mean analysis.

To address the problem of extreme outliers, a new technique is developed that randomly divides the full ensemble into sub-ensembles of equal size for each observation at each analysis step. All sub-ensembles are updated independently using deterministic filter algebra. The random subgrouping technique removes the effects of extreme outliers in two ways: the smaller ensemble size for each sub-ensemble limits the deviation of an outlier from the ensemble mean and the random subgrouping prevents the long-term persistence of an outlier.

Test results, using the random subgrouping technique on two low-order models (Lorenz-63 and Lorenz-96) and a global QG atmospheric model coupled to a slab ocean, show that the new scheme significantly improves performance compared to regular stochastic and deterministic filters.

2.1. Introduction

First introduced by Evensen (1994), the ensemble based filter is emerging as a powerful tool for data assimilation (Evensen 2007). The key element of the filter is to derive the forecast uncertainties from an ensemble of model integrations. Ensemble based filters can be divided into two types: stochastic and deterministic. The two types of methods differ mainly in how they update the ensemble to match the analysis uncertainty. The updated analysis variance is produced by the error variances from both the forecast and observation. A deterministic filter, also called an Ensemble Square Root Filter (EnSRF) (Anderson 2001, 2003, Bishop et al. 2001, Whitaker and Hamill 2002, Tippett et al 2003, Sakov and Oke 2007), transforms the ensemble anomaly to match the variance given by the Kalman Filter (KF) theory (Kalman 1960). In contrast, a stochastic filter (EnKF, in this chapter) attempts to match the updated variance from the KF theory by adding perturbations to the observations (Burger et al. 1998, Houtekamer and Mitchell 1998).

The ensemble based filter uses the ensemble spread to represent the forecast and analysis uncertainties with the assumption of Gaussian white. This method introduces two error sources into the analysis: the sampling errors from the limited ensemble size (Whitaker and Hamill 2002, Sacher and Bartello 2008) and the non-Gaussian probability density function (PDF) of the error from the nonlinear system. The sampling error in EnKF appears in both the background uncertainty and observational uncertainty. An EnSRF avoids the sampling error introduced by perturbed observations and tends to generate better analyses than an EnKF when applied to a linear model, especially for a small ensemble size (~10-20) (Whitaker and Hamill 2002, Evensen 2003, Anderson 2010).

An EnKF performs better than an EnSRF when applied to a nonlinear system with non-

Gaussian PDFs for a large ensemble size because the accuracy of EnSRF is severely compromised by the effect of extreme outliers (Lawson and Hansen 2004, Anderson 2010). An outlier is a member that deviates markedly from the general population of ensemble members. A strongly nonlinear system produces outliers in an ensemble. An EnSRF generates persistent outliers because it has no effective way to restore the outliers to the majority, thus leading to extreme outliers and a large analysis error. The outliers in EnKF are relatively less extreme because they can be mixed with the general ensemble members through randomly perturbed observations. Furthermore, a larger ensemble size is more likely to produce a stronger outlier effect, so an EnSRF applied to a nonlinear system will perform worse as the ensemble size increases (Lawson and Hansen 2004, Mitchell and Houtekamer 2009, Anderson 2010).

Several methods have recently been proposed to improve the performance of the EnSRF in eliminating the effect of extreme outliers in nonlinear systems. Sakov and Oke (2008) use a random transformation to decrease the impact of extreme outliers in the ensemble transform filter, while Anderson (2010) uses a rank histogram filter to eliminate extreme outliers. In this study a new filter scheme, called the random subgrouping EnSRF, is proposed to eliminate the distortion effect of extreme outliers and therefore to improve the filter performance in a nonlinear system. The random subgrouping EnSRF divides the entire ensemble randomly into subgroups of equal size and updates each subgroup independently using EnSRF. In comparison to current EnSRF and EnKF, the new scheme significantly improves the filter analysis in tests of several nonlinear systems. In section 2, the algorithm of the ensemble based filter and random subgrouping scheme are briefly described. In section 3, we will demonstrate the performance of the random subgrouping scheme, first using two simple “toy” models and then using an intermediate model. A summary will be given in section 4.

2.2 The Algorithm of Ensemble Based Filters

The ensemble based filter is a Monte Carlo implementation of the KF. It represents the distribution of the system state using an ensemble) and replaces the covariance matrix by the sample covariance computed from the ensemble. A successful ensemble-based filter need to achieve an analysis ensemble representing the real uncertainty of analysis. Based on the methods of updating analysis ensemble, ensemble based filters can be divided into two types: stochastic and deterministic.

2.2.1 Stochastic filter (EnKF)

A stochastic filter (EnKF) treats each observation as a random variable that is perturbed to sample the observational uncertainty when it updates the ensemble. So the eqn. (1.14) is written as

$$\mathbf{x}_i^a = \mathbf{x}_i^f + \mathbf{K}_e(\mathbf{y}_i^o - \mathbf{H}\mathbf{x}_i^f) \quad (2.1)$$

where the ensemble index $i = 1, 2, \dots, N$ and N is the ensemble size. The resulting analysis ensemble \mathbf{x}_i^a has a mean and variance that matches the theoretical values in equations of (1.14) and (1.15) in chapter 1 (Burgers et al 1998, Houtekamer and Mitchell 1998). To ensure no extra noise is added by the perturbations, the ensemble mean perturbation is set equal to zero (Pham 2001; Mitchell et al. 2002, Evensen 2007, Mitchell and Houtekamer 2009). A flow chart of an analysis cycle for EnKF is shown in Fig.2.1

2.2.2 Deterministic filter (EnSRF)

A deterministic filter (EnSRF) transforms the forecast ensemble to match the analysis and

its uncertainty (see eqn. (1.11) and (1.12) in chapter 1).

This chapter will show the results from one particular EnSRF, the Ensemble adjustment filter (EAKF, Anderson 2001, 2003). EAKF updates the ensemble in two steps (Anderson 2003). First it derives the analysis ensemble mean and variance and then computes the ensemble increment to match the analysis error in the observation space. In the second step, the ensemble increment is distributed over relevant state variables using a least square fitting.

2.2.3 The Random Subgroup Ensemble Based Filters

An ensemble based filter can produce extreme outliers in nonlinear systems. This phenomenon is particularly notable in using EnSRFs with the simple nonlinear models (Lawson and Hanson 2004, Anderson 2010). In this study, a new filter scheme, called random subgrouping EnSRF is developed to eliminate the effect of outliers that often occur in a regular EnSRF.

This new scheme is the same as a regular EnSRF except that for each observation at each analysis step, the entire ensemble is divided randomly into sub-ensembles of an equal size. All the sub-ensembles are updated independently using the same observation, but with their own forecast covariance. Each sub-ensemble will have a different combination of ensemble members for different observations at different analysis times. The final analyses and forecasts are constructed using the entire ensemble. The difference between a sEnSRF and an EnSRF or an EnKF in an analysis step can be seen in Fig.2.1. A random subgrouping EnSRF cycle is performed in 4 steps:

- (1) the model ensemble is integrated forward to the next observation time;
- (2) the N -member ensemble is randomly divided into n sub-ensembles of equal N/n size;

- (3) each sub-ensemble is updated independently using EnSRF with the same observation;
- (4) steps 2 to 3 are repeated until the all the observations are used.

The new scheme removes the effect of extreme outliers in two ways. The smaller ensemble size for each sub-ensemble limits the magnitude of deviation of an outlier from the ensemble mean; and the random subgrouping prevents an outlier from being persistent.

The random subgrouping EnSRF shares the same algebra as EnSRF, so the sampling error generated by the perturbed observations in EnKF is avoided. It is still a square root method but no longer deterministic since random subgrouping introduces the stochastic variability into the filter system. The objective of our sub-grouping procedure differs from that of Houtekamer and Mitchell (1998), where they use a sub-grouping process to reduce the negative bias in the analysis error variance. The negative bias comes from the use of the same ensemble to estimate both by P_e^f and K_e in equations of (1.14) and (1.15) in chapter one.

Following (Lawson and Hansen 2004 and Anderson 2010), the ensemble kurtosis

$$Kurtosis = \frac{\frac{1}{N-1} \sum_{i=1}^N [x_i - E(x_i)]^4}{\sigma^4} \quad (2.2)$$

is used to quantify the presence of outliers, where σ^2 is the sampling variance and $E(x_i)$ is ensemble mean. The ensemble kurtosis increases with increasing ensemble size for both EAKF and EnKF, (for example for Lorenz-63 system, Fig.11 of Anderson 2010). An EnSRF conserves higher-order moments through the analysis step (Anderson 2001). As a result, the presence of extreme outliers and a corresponding large ensemble kurtosis must occur during the forecast steps.

An extreme outlier will deviate less from the ensemble mean for smaller size ensembles. Let us consider a one-dimension case in observational space, where we define a forecast

uncertainty of σ_b^2 and the perturbation of each regular ensemble member from the ensemble mean to be y_i , ($i = 1, 2, 3, \dots, N$), and N to be the ensemble size. For convenience, the N^{th} member is set as an outlier and its perturbation from the ensemble mean is denoted as y_N . We then have

$$y_N + \sum_{i=1}^{N-1} y_i = 0 \quad (2.3)$$

$$y_N^2 + \sum_{i=1}^{N-1} y_i^2 = (N-1)\sigma_b^2 \quad (2.4)$$

For $y_i = y^m + \varepsilon_i$, and $\sum_{i=1}^{N-1} \varepsilon_i = 0$, where $i = 1, 2, 3, \dots, N-1$, and y^m , the ensemble mean excluding the outlier, we have

$$y^m = -\frac{1}{N-1}y_N \quad (2.5)$$

$$y_N^2 + \frac{1}{N-1}y_N^2 + \sum_{i=1}^{N-1} \varepsilon_i^2 = (N-1)\sigma_b^2 \quad (2.6)$$

$$y_N^2 \leq \frac{(N-1)^2}{N}\sigma_b^2 \quad (2.7)$$

$$|y_N| \leq \frac{N-1}{\sqrt{N}}\sigma_b \approx \sqrt{N}\sigma_b \quad (2.8)$$

For a given forecast uncertainty σ_b^2 , which is constrained by the observational uncertainty, the maximum deviation $|y_N|$ for an extreme outlier is therefore roughly proportional to the square root of the ensemble size N . When $N=100$, the maximum deviation can reach to $9.9\sigma_b$. For $N=5$, the deviation is smaller than $2\sigma_b$, which is comparable to the deviation of a regular ensemble member. When the outlier reaches its maximum deviation, the ensemble kurtosis also reaches the maximum its of $\frac{(N-1)^3+1}{N^2}$. Therefore the outliers are potentially much severe for a large ensemble than for a small ensemble. The similar derivation works for analysis ensemble with an uncertainty of σ_a^2 .

When we use the subgrouping scheme to update the ensemble, the subgroups without the outlier have small σ_b and consequently the Kalman gain and the correction towards the

observation will be small; the subgroup with the outlier will have large σ_b and consequently the correction towards the observation can be large. The subgrouping thus acts to selectively pull the outlying member(s) towards the other members. With the subgrouping scheme, the total outlier deviation $|y_N|$ is the composite of the deviation from its own sub-ensemble mean and the deviation of its sub-ensemble mean from the full ensemble mean. The first part is bounded by the small sub-ensemble size (eqn. 2.8). However the second part can contribute significant deviation for a subgrouping scheme without randomization. Without randomization the sub-ensembles are independent from each other. The sub-ensemble mean may significantly diverge when the sub-ensemble size is small (5~10). With randomization each sub-ensemble has a different combination for different observations and different times. In that way each ensemble member will benefit from all the other ensemble members. Thus, the deviations of sub-ensemble mean from the full ensemble mean are small and the total outlier deviation $|y_N|$ is small.

The randomization during subgrouping process further decreases the effect of extreme outliers by preventing their persistence. When an extreme outlier happened in a subgroup, the deviation of the outlier is much bigger than the deviation of other members. This relation can propagate to the next analysis cycle through the forecast step. Therefore the outlier tends to be persistent at the following analysis cycles. A persistent outlier will contribute to the development of a severe deviation. The relation will be break by the randomization during subgrouping process. A dominant extreme outlier in one subgroup combination, however, may not be dominant in another subgroup combination. The random subgroup scheme gives different combinations for subgroups, and therefore prevents the persistence of outliers.

The random subgrouping scheme introduces sampling errors into the filter simulation. Each sub-ensemble has a much smaller sample size than the full ensemble. The estimated

background uncertainty for an individual sub-ensemble only uses part of the information of the full ensemble, and therefore has greater sampling error than that is estimated from the full ensemble.

The sub-ensembles are independent of each other for subgrouping without randomization. The random subgrouping scheme connects whole ensemble together, and each ensemble member still benefits from all of the peers. We suspect that the sampling error for the random subgrouping scheme is less than for subgrouping without randomization.

In summary, a random subgrouping scheme reduces the effect of extreme outliers but introduces more sampling error. It improves the filter performance in a non-linear system when the outlier effect is strong. However, the analysis error could increase when the system is linear or the outlier effect is negligible. In the next section, we will show that the random subgrouping EnSRF performs better than EnKF and EnSRF in tests with several nonlinear systems.

2.3 Results

Here the random subgrouping scheme with a specific EnSRF – EAKF is applied to three nonlinear systems: two simple models (Lorenz-63, Lorenz-96) and one inter-median model (QG-slab). These models are described in the appendices. A random subgrouping EAKF with a sub-ensemble size of n will be denoted as $sEAKF_n$. Our $sEAKF_n$ results will be compared with the results from EnKF and EAKF applied to the same systems using identical observations and initial conditions.

2.3.1 $sEAKF_n$ in the Lorenz-63 model

Firstly we apply $sEAKF_n$ to the Lorenz-63 model that has three variables (x, y, z) . For all

experiments except in section (D) we apply an idealized observation system which covers all three variables with an observation frequency of 0.1 time unit (10 time steps). The details of model and experimental design are shown in appendix 2A.

(A) Large ensemble experiments

We first examine large ensemble cases, specifically, an ensemble size of 80 subdivided into 16 subgroups each with a sub-ensemble size of 5 (sEAKF₅). Due to the strong nonlinear nature of the Lorenz-63 model, outliers are produced in the 80-member ensemble for the EnKF and EAKF simulations. However, no outlier is produced in the sEAKF₅ simulation (Fig. 2.2).

In the EAKF case (Fig.2.2b), an extreme outlier persists for more than 20 time units (200 analysis cycles) with a deviation far from the other ensemble members. This deviation occurs because in EAKF there is no mechanism to prevent it from drifting from the other ensemble members. An EAKF tends to retain high-order moments through the assimilation process (Anderson 2001), thus leading to the outliers that can persist during the EAKF assimilation cycles. The persistence of outliers produces big deviation from most other members, which are constrained to be in close proximity to one another to balance the outlier deviation (Fig. 2.2b).

In EnKF, there are also extreme outliers deviating significantly from the majority of ensemble members (Fig. 2.2a), but their magnitude and persistence are much smaller than that for EAKF (Fig.2.2b). This is because random perturbations in the observations play the role of adding noise to outliers in EnKF, which works to restore an extreme outlier toward the main population after a few cycles of analysis. In comparison, our random subgrouping scheme (sEAKF₅) essentially eliminates all the extreme outliers (Fig. 2.2c).

The time-mean ensemble kurtosis of the Lorenz y -variable is ~ 19 for the EAKF with an 80-member ensemble. This is much larger than the kurtosis value of 3 for a Gaussian distribution. The time-mean ensemble kurtosis of y -variable is ~ 4 for EnKF with the 80-member ensemble, which is much less than for EAKF and a little greater than Gaussian. The kurtosis of y -variable in sEAKF₅ is ~ 2.5 , which is even smaller than the kurtosis for the Gaussian distribution. This quantitatively verifies that few spurious outliers are generated by this new assimilation scheme. Therefore, sEAKF₅ has a reduced outlier effect relative to EnKF and, especially, EAKF.

The ensemble spread represents the uncertainty (error) of the ensemble mean in an ensemble-based filter. We expect that ensemble spread and analysis root mean square error (RMSE) are consistent during filter simulation. The sEAKF₅ and EnKF simulations show statistical consistency between the values of RMSE and ensemble spread (Fig.2.3a). The ensemble spreads are slightly bigger than the RMSE for sEAKF₅ simulations and slightly smaller for EnKF simulations (Fig.2.3a – blue and green squares). EAKF produces a large difference between RMSE and ensemble spread because persistent outliers distort the PDFs of forecast and analysis. Some EAKF experiments produce very big analysis errors (with the RMSE > 0.9) but very small ensemble spreads (~ 0.7). On average this leads to greater RMSE (0.84) than ensemble spread (0.66) (Fig. 2.3a – red square).

The sEAKF₅ generates the smallest analysis error among the three filter schemes. The analysis RMSE for sEAKF₅ is 0.57 that is significantly smaller (with 99% confidence) than that for EAKF (0.84) or for EnKF (0.61) (Fig. 2.3b). Starting from identical first guesses (initial conditions) and using identical observations, 95% (85%) of sEAKF₅ experiments produce smaller RMSE than corresponding EAKF (EnKF) experiments (Fig. 2.3b).

(B) Small ensemble experiments

The sEAKF_n scheme also improves filter performance on the Lorenz-63 system for smaller ensemble sizes (<80). EAKF simulations with a 20-member ensemble, which have an average ensemble kurtosis of about 3.6, produce some extreme outliers but less than for the large ensemble size (80) (Fig. 2.4). The performance of the 20-member EAKF (Fig. 2.4) is better than the performance of the 80-member EAKF (Fig. 2.3) because of the less extreme outlier effect, which is consistent with previous studies (Lawsen and Hansen 2004, Mitchell and Houtekamer 2009, Anderson 2010). sEAKF₅ simulations (4 subgroups) with a 20-member ensemble produce smaller RMSE than the corresponding EAKF simulations (0.58 vs. 0.63) (with 99% confidence) (Fig. 2.4). Because of the small ensemble size, the sampling error caused by the perturbed observations in EnKF simulations exceeds the outlier effect in the EAKF simulations. The EnKF simulations perform poorest among the three schemes (Fig. 2.4).

The small ensemble size also leads to an ensemble spread for EnKF that is smaller than RMSE (Fig. 2.4a). The sEAKF₅ and EAKF show more consistency between ensemble spread and RMSE than occurs in EnKF. As in the large ensemble case, the ensemble spreads are slightly bigger than the RMSE for sEAKF₅ simulations and slightly smaller for EAKF simulations. For identical first guesses (initial conditions) and observations, more than 80% of sEAKF₅ experiments produce smaller RMSE than the corresponding EAKF and EnKF experiments (Fig. 2.4b).

In summary, for both large and small ensemble experiments, sEAKF₅ significantly improves the data assimilation quality in the Lorenz-63 system compared with EnKF and EAKF. Additionally, the slightly bigger ensemble spreads compared to RMSEs in both large and small

ensemble experiments indicate that our sEAKF₅ assimilations do not suffer the negative bias of analysis ensemble spread which usually happens in the regular ensemble based filters (first pointed out by Houtekamer and Mitchell 1998). The exact reason for this needs to be explored in further research.

(C) Optimal size of the sub-ensemble

The random subgrouping scheme uses a small sub-ensemble size to eliminate the effect of extreme outliers but this procedure also introduces more sampling error into the filter system. More precisely, smaller sub-ensemble sizes lead to a less extreme outlier effect but bigger sampling errors.

An important consideration is determining a sub-ensemble size that optimally balances the effect of extreme outliers and sampling error. It would be useful to balance the effects of extreme outliers with the sampling error to obtain an optimal sub-ensemble size. On the one hand, the sub-ensemble size cannot be too big because it will have extreme outliers. For example, the ensemble kurtosis of the y-variable is ~ 4.6 for sEAKF₄₀ (2 subgroups) in the 80-member ensemble experiments and ~ 9.2 for sEAKF₈₀ in 160-member ensemble experiments (Fig. 2.5a). On the other hand, if the sample size for the sub-ensembles is too small, the sampling error becomes too large. For a subgroup of only 2 members (sEAKF₂), the ensemble kurtosis is in a reasonably small (~ 2.8) (Fig. 2.5a) but the analysis RMSE is large (~ 0.7) (Fig. 2.5b)

The sub-ensemble size limits the magnitude of outlier deviation from the ensemble mean. A larger ensemble size increases the deviation for outliers and therefore requires more subgroups to eliminate them. As a result, for various ensemble sizes (20, 40, 80 and 160) the ensemble kurtosis of y-variable converges to a minimum value of 2.5 for a sub-ensemble size of 5 (Fig.

2.5a). Hence the minimum kurtosis seems to be achieved at a size of sub-ensemble independent of full ensemble size.

A smaller ensemble kurtosis implies a smaller effect of the outliers and therefore an improved filter performance (see RMSE – Fig. 2.5b). For example, for the 80-member ensemble size, sEAKF₅ gives the smallest ensemble kurtosis as well as analysis RMSE, which is reduced by ~30% from that of the standard EAKF.

The optimal sub-ensemble size may change when the experimental design changes, such as the nature of observation frequency and uncertainty. However, optimal sub-ensemble size cannot be smaller than 5 because extreme outliers do not occur for this smaller ensemble size (Eqn. 2.8). It is also worth noting, however, that even using only 2 subgroups, sEAKF_n can significantly decrease ensemble kurtosis and analysis RMSE compared to EAKF in Lorenz-63 system. For example, the ensemble kurtosis and analysis RMSE (3.0 and 0.59) for sEAKF₂₀ with full ensemble size 40 are much smaller than those in corresponding EAKF simulations (6.5 and 0.66) (Fig. 2.5).

The effect of the randomization process during subgrouping can be seen explicitly by comparing the results from subgrouping experiments with and without randomization. Consistent with our discussion in section 2 (C), the subgrouping simulations without randomization give a bigger ensemble kurtosis and analysis RMSE than the corresponding simulations with randomization (sEAKF_n) (Fig. 2.5).

(D) Experiments with different observation frequency

We further tested the sEAKF_n on the Lorenz-63 system by using different observation frequencies (Fig. 2.6). The optimal sub-sample size varies with the change of observation

frequency. For example, it is 20 for an observation frequency of 0.01 and 10 for an observation frequency range from 0.02 to 0.05 (figure not shown).

sEAKF_n with an optimal sub-ensemble size performs better than the corresponding EAKF because it suppresses the extreme outlier effect. However, the improvement of sEAKF_n over EAKF is relatively small when the observation time interval is small (0.01~0.05) and the extreme outlier effect in EAKF is weak (see the kurtosis in Fig. 2.6a). The biggest improvement in filter performance of sEAKF_n over EAKF occurs when the observation time interval is 0.2 (20 time steps) and the EAKF ensemble kurtosis of y-variable is greatest (25.7). The analysis RMSE of sEAKF₅ is only 60% of the RMSE from the corresponding EAKF simulations.

The sEAKF_n also performs better than EnKF in all the observation frequency experiments (Fig. 2.6b). The ensemble kurtosis of y-variable for all EnKF simulations with an 80-member ensemble in Fig.2.6a is relatively small (around 3.2~4), which indicates a very weak outlier effect in the EnKF simulations. The corresponding ensemble kurtosis for all sEAKF_n simulations is ~2.5, which is consistent with results shown in Fig. 2.5a.

It is interesting to note that the ensemble kurtosis of y-variable for the assimilations with an optimal sub-ensemble size do not converge to the kurtosis of a Gaussian distribution (3), but rather to a smaller value (~2.5). While the reason for this is unclear, it may be related to the PDF of Lorenz-63 system. Since the system kurtosis for the y-variable of the Lorenz-63 system (derived from 10^4 time units of control integration) has a value of ~2.5.

(E) Ensemble forecast

A good analysis produces a good initial condition for the forecast. We also examine the ensemble forecasts by using the analyses of 80-member ensemble simulations as initial

conditions. The ensemble forecasts are called EnKF forecast, EAKF forecast and sEAKF₅ forecast separately based on the initial condition sources. The details of forecast experimental design are shown in appendix A. The results show that the sEAKF₅ provides the best initial conditions for the ensemble forecast among the three filter schemes (Fig. 2.7).

The RMSE of sEAKF₅ forecast is significantly smaller than that of EAKF forecast (Fig. 2.7a). For example, the average forecast RMSEs at time 0.5 are 2.5 for sEAKF₅ forecast and 3.5 for EAKF forecast. The RMSE of EnKF forecast is slightly bigger than that of sEAKF₅ forecast. The forecast ensemble spread indicates the forecast uncertainty of the ensemble forecast. The sEAKF₅ and EnKF forecasts show statistical consistency between the values of RMSE and ensemble spread from forecast time 0-0.5 and then the ensemble spreads become slightly bigger than the RMSEs (Fig. 2.7a). The ensemble spread of EAKF forecast shows consistent, significant negative bias compared to its RMSE.

The skill-spread correlation, calculated as the correlation between the forecast RMSE and the ensemble spread for the total 20000 forecasts, is another important measure of the quality of a forecast scheme (Fraedrich and Ziehmann-Schlumbohm, 1994; Anderson, 1997; Hamill et al., 2000). A high correlation implies a higher likelihood of prediction of the forecast accuracy from the forecast spread. Fig.2.7b shows that sEAKF₅ forecast has significantly higher skill-spread correlation over the EnKF forecast and especially over EAKF forecast.

One may be concerned that the relatively low value of analysis kurtosis of sEAKF₅ may be too small for sEAKF₅ forecast to provide a good probabilistic prediction of extreme events, which is important for an ensemble forecast. That is not a problem in our forecast experiments because the subgrouping scheme only limits the ensemble kurtosis of analysis and the kurtosis increases during the forecast. Fig. 2.7c shows that the ensemble kurtosis increases from ~2.5 to

~ 4.5 in sEAKF₅ forecast and from ~ 4 to ~ 9.5 in EnKF forecast after 1.0 forecast time has elapsed. The increasing trend of kurtosis does not exist in EAKF forecast because its kurtosis value is already extraordinary high at the beginning. We use three times the ensemble spread as a threshold, which accounts for 99.7% of the sample population, assuming the ensemble distribution is normal. A forecast will be regarded as forecast failure (missing the extreme event) when its RMSE is greater than the threshold. The probabilities of forecast failure are around 0.2-0.5% in sEAKF₅ forecast and around 0.5-0.8% in EnKF forecast. These probabilities are much smaller than that in EAKF forecast (1.7-4.6%) (Fig. 2.7d).

Therefore, the sEAKF₅ forecast provides the best forecast among the three schemes without sacrificing the predictions of extreme events.

(F) Subgrouping in EnKF

Because there are weak effects from outliers in EnKF simulations with big ensembles, EnKF can also benefit from random subgrouping. The kurtosis of EnKF is ~ 4 for 80-member ensemble experiments and ~ 5.5 for 160-member ensemble experiments (Fig. 2.5a), which indicate outliers occurring during assimilation. One can use the random subgrouping to eliminate the weak outlier effect and improve the filter performance. The steps of applying random subgrouping to EnKF are the same as for EnSRF except that the EnKF algebra is used in step 3 when the observations are assimilated sequentially.

We use the Lorenz-63 model to test the random subgrouping scheme in EnKF assimilations (Fig. 2.8). The random subgrouping EnKF with a sub-ensemble size of n is indicated as sEnKF _{n} . For all three experiment settings (ensemble size of 40, 80, 160), the sEnKF _{n} s with an optimal sub-ensemble size perform significantly better than the corresponding

regular EnKFs (see the smaller RMSE in Fig.2.58b). Similar to the sEAKF_n, the optimal sub-ensemble size of sEnKF_n is also independent of full ensemble size. The optimal sub-ensemble size is 10 in our experiments (Fig.2.8b). When the sample size for the sub-ensembles is too small and the sampling error becomes larger than the outlier effect, the sEnKF_n(s) generate larger RMSE than the corresponding EnKFs. The average ensemble kurtoses in the sEnKF_n simulations converge to the value of ~ 3 (Fig.2.8a), which is smaller than the average ensemble kurtosis in regular EnKFs.

The Kalman gain for a sub-ensemble in our EnKF assimilation with random subgrouping is computed from all of the members in the sub-ensemble. This is different from the subgrouping approach of Houtekamer and Mitchell (1998) where each sub-ensemble uses the Kalman gain computed from all of the other members out of the sub-ensemble.

We can also apply randomization to the Houtekamer and Mitchell's subgrouping scheme. The testing results on the Lorenz-63 model, however, do not show significant difference between experiments with and without randomization. In summary, Houtekamer and Mitchell's subgrouping scheme produces better consistency between RMSE and ensemble spread but bigger analysis RMSE compared with our random subgrouping EnKF (not show).

2.3.2 sEAKF_n in Lorenz-96 model – model independence

(A) Different influence radius

When a system includes more spatial dimensions and more variables (compared to say the Lorenz-63 model), the covariance between different variables or different locations becomes important and may lead to amplification of the sampling error introduced by subgrouping. To address this potential amplifying effect, sEAKF_n is applied to another simple nonlinear system -

the Lorenz-96 model with 40 dimensions (see Appendix 2B for detail). Since this idealized model is more complex than the Lorenz-63 system, the Gaspari and Cohn (1999) covariance localization is used during assimilation. The covariance localization influence radii are denoted as D . The details of experimental design are shown in appendix 2B.

The RMSEs for sEAKF_n, EAKF and EnKF with different values of D are shown in table 2.1. Based on RMSE, there is a preferred D for a given sub-ensemble size (n) for sEAKF_n. For example, the preferred D is 7 for sEAKF₅ and 9 for sEAKF₁₀. The relationship of smaller preferred D with smaller sub-ensemble size (n) for sEAKF_n might identify the amplifying effect in our experiments. However, this amplifying effect turns out to be negligible for sEAKF_n in Lorenz-96 system. This is because sEAKF_n, with an optimal sub-ensemble size (20) and a range of influence radii (D) produces a smaller RMSE than occurs with EAKF and EnKF. Both EAKF and sEAKF₂₀ reach their minimum RMSE when $D = 11$.

(B) Different levels of chaos

The effectiveness of sEAKF_n for systems of different levels of nonlinearity is also examined by using different forcing (F) in the Lorenz-96 model (table 2.2). Here the influence radius D is set at 7. When the system is periodic ($F=2$), the outlier effect is minor for an 80-member ensemble simulation. The RMSEs for EAKF, sEAKF₄₀ and sEAKF₂₀ are comparable, implying insensitivity to the subgrouping scheme. As the system becomes chaotic ($F = 5, 8, 10$), the sEAKF_n simulations with an 80-member ensemble perform significantly better (with 99% confidence) than the corresponding EAKF and EnKF simulations. The RMSE from sEAKF_n with optimal sub-ensemble size (10 for $F=5$ and 20 for $F=8, 10$) is reduced by 8~9% (7~12%) from that of the EAKF (EnKF).

2.3.3 sEnSRF in QG-slab model

The improvement in performance of sEAKF_n is also obtained in tests with a more realistic model – a global quasi-geostrophic atmosphere model coupled with a slab ocean (QG-slab model) (details in Appendix C). The globally-averaged RMSE of the streamfunction (ψ) for sEAKF₄₀ (two subgroups for an 80-member ensemble) in our experiments is $7.65 \times 10^4 m^2 s^{-1}$, which is significantly smaller (with 99% confidence) than that for EAKF ($7.95 \times 10^4 m^2 s^{-1}$) and for EnKF ($8.15 \times 10^4 m^2 s^{-1}$).

We speculate that the improved performance of sEAKF₄₀ in the QG-slab model is due to the removal of the extreme outliers by the random subgrouping scheme. The extreme outlier effect usually occurs in EnSRF with a large ensemble (Amezcuca et al 2012), while the ensemble size of 80 in our experiments is much smaller than the model grid size (64 longitude \times 54 latitude grid points). The smaller RMSE for EAKF, compared to that for EnKF, indicates that the outlier effect in the QG-slab model is not as severe as it is in simple models like Lorenz-63 and Lorenz-96. Furthermore, there are no permanent outliers because the extreme outlier in EnSRF is a transient phenomenon (Amezcuca et al 2012). However, the reasons above do not prevent extreme outliers from occurring in some local regions at certain times. For example, large kurtosis develops in the southern Indian Ocean and the Caribbean and persists during analysis cycles of 569-572 in one EAKF experiment (Fig.2.9a). The large kurtosis in the Caribbean is more stationary, while the large kurtoses in the southern Indian Ocean propagate westward. The persistent large kurtosis phenomenon is not found in the same assimilation period in the corresponding sEAKF₄₀ experiment, which uses the same initial conditions and observations as the EAKF experiment (Fig.2.9b). Averaged from all of the EAKF experiments, 0.30% of the ensemble kurtoses are larger than 5, which is significantly larger than the percentages in

sEAKF₄₀ assimilations (0.15%). Therefore, the outlier effect, though weak, does exist in our 80-member EAKF assimilation. Although it is not a big drawback for EAKF performance, we can still alleviate the outlier effect using the random subgrouping scheme to further improve the filter performance.

That the random subgrouping scheme improves the filter performance by alleviating the outlier effect can also be seen in the RMSE consistency between EnKF and sEAKF₄₀. If an extreme outlier effect in EAKF is significant in a certain region, we expect that both EnKF and sEAKF_n will also perform better than EAKF in that region because the outlier effect is trivial in the EnKF and sEAKF_n assimilations. In each set of experiments, which used identical initial guesses and observations, the analysis RMSE difference between EAKF with and without random subgrouping (EAKF - sEAKF₄₀) (Fig.2.10a) shows a good spatial consistency with the RMSE difference between EAKF and EnKF (EAKF - EnKF) (Fig.2.10b). This consistency is very robust, although the spatial patterns of the analysis RMSE difference vary among different sets of experiments. The average spatial correlation coefficient from 32 sets of experiments is around 0.4, which is above the 99% confidence level. In each set of experiments, the RMSE of sEAKF₄₀ is not uniformly smaller than that of EAKF (Fig.2.10a) because the extreme outlier in EAKF is only a transient phenomenon (Amezcuca et al 2012). Averaging all 32 sets of experiments, the sEAKF₄₀ does perform significantly better than the corresponding EAKF in most regions (see the RMSE difference in Fig.2.10c)

Therefore, we speculate that EAKF in QG-slab model with an ensemble size of 80 does suffer from a weak effect of extreme outliers and the better performance of sEAKF₄₀ results from reducing that outlier effect.

In summary, the improvement in assimilation results of sEAKF_n over EAKF and EnKF

for different models (Lorenz-63, Lorenz-96 and QG-slab) suggests that our filter scheme is a robust method that should be tested on complex weather and climate models.

2.4. Summary

The extreme outlier problem arising from non-Gaussian PDFs is a challenge for an ensemble based filter, this is especially true for EnSRF using a large ensemble size (Lawsen and Hansen 2004, Anderson 2010, Lei et al. 2010). To address the problem of extreme outliers, a new technique is developed that randomly divides the full ensemble into sub-ensembles of equal size for each observation at each analysis step. All sub-ensembles are updated independently using deterministic filter algebra. The random subgrouping technique eliminates outliers in two ways. First, the smaller size for each sub-ensemble limits the deviation of an outlier from the ensemble mean. Second, the randomization process prevents the long-term persistence of an outlier.

The random subgrouping EnSRF shares the same algebra as EnSRF, so the sampling error generated by the perturbed observations in EnKF is avoided. It is still a square root method but no longer deterministic because it uses a randomization process in creating subgroup ensembles. In general, our random subgroup scheme shares the same principle as the random transformation method in Sakov and Oke (2008): alleviate the outlier effect by adding some randomization into EnSRF system. Our random subgrouping EnSRF introduces additional sampling errors by subgrouping, while the random transformation method loses the traceability of individual ensemble trajectories (Amezcuca et al 2012).

As a test, the random subgrouping scheme with a specific EnSRF – EAKF (sEAKF_n) is applied to three nonlinear systems: two simple models (Lorenz-63, Lorenz-96) and one inter-

median model (QG-slab). The new scheme significantly improves the filter analysis relative to both the stochastic filter (EnKF) and the deterministic filter (EnSRF). It is worth pointing out that the outlier effect, though weak, exists in our QG-slab model EAKF assimilation, even if the ensemble size is much smaller than the model state dimension. We speculate that the weak outlier effect could exist in some geophysical EnSRF assimilations. The random subgrouping scheme can alleviate the weak outlier effect to further improve the performance of EnSRF.

In terms of future potential, the random subgrouping scheme can be easily incorporated into any of the current deterministic filter systems. Our tests with the simple models and intermediate models suggest that this new filter scheme will be particularly effective in suppressing extreme outliers in highly chaotic systems. We propose further testing of the random subgrouping scheme in ensemble based filter with more comprehensive models. .

2.5 Appendix 2A: Lorenz-63 model

The Lorenz-63 model (Lorenz 1963) describes one of the most famous nonlinear dynamical system, is described in the equations,

$$\dot{x} = \sigma(y - x) \quad (A2.1)$$

$$\dot{y} = \beta x - y - xz \quad (A2.2)$$

$$\dot{z} = xy - cz \quad (A2.3)$$

The parameter β is the ratio of the Rayleigh number divided by the critical Rayleigh number. The parameter σ is the Prandtl number. The third parameter c is related to the horizontal wave number of the system. By choosing typical values of the parameters ($\beta = 28$, $\sigma = 10$, $c=8/3$), the evolution of the state vector (x, y, z) follows the well-known Lorenz attractor.

The model is integrated using a 4-th order Runge-Kutta method with a time resolution of

$dt = 0.01$ (~1 hours if we treat one time unit as 4 days). We first generate the “truth” in a long control simulation of 10^4 time units.

For each set of experiments, we first generate the “observations” by adding onto the “truth” random errors with a standard deviation (2, 2, 2). The initial first guess is then randomly picked from the “observations”. The initial ensemble is built based on the initial first guess and its uncertainty (error scale). Using the initial ensemble and “observation”, the model is simulated for 500 analysis cycles by using different data assimilation schemes (SEAKF_n, EAKF, EnKF) with a given observation time interval. The observation time interval is 0.1 for all experiments except Fig.2.6. Following a spinup of 200 analysis cycles, the results from all analysis steps are used to calculate the RMSE, kurtosis, ensemble spread for the experiment.

500 sets of experiments are performed. For a fair comparison we did not use any inflation scheme in all the experiments.

The analyses of 80-member ensembles, with observation time interval of 0.1, are used as initial conditions to test the ensemble forecast. We randomly pick 40 initial conditions from each set of experiments for different data assimilation schemes (SEAKF₅, EAKF, EnKF). Based on the initial condition sources, the ensemble forecasts are called EnKF forecast, EAKF forecast and sEAKF₅ forecast separately. Each forecast has 20,000 experiments with forecast time of 1.0 (100 time steps).

2.6 Appendix 2B : Lorenz-96 model

The Lorenz-96 model is a latitude circle model first proposed by Lorenz (1996) to study fundamental issues regarding the forecasting of spatially extended chaotic systems such as the atmosphere. It has N state variables governed by equation

$$\dot{X}_i = (X_{i+1} - X_{i-2})X_{i-1} - X_i + F \quad (A2.4)$$

where $i = 1, \dots, N$ are the cyclic indices. We use $N = 40$ for our simulations. The model is integrated using a 4-th order Runge-Kutta method with a time resolution of $dt=0.005$. To investigate the filter performance under different conditions, we assign values (2, 5, 8, 10) for the forcing term F that defines the system shifting from periodic (2) to chaotic (5, 8, 10) (Lorenz 2005).

The experiment design is similar to Appendix 2A but all the simulations use an ensemble size 80. The observations cover all the grids with a frequency of 0.1 (20 time steps) and the observation error standard deviation of 2. The Gaspari and Cohn (1999) covariance localization with different influence radius D is used in all the simulations (Houtekamer and Mitchell, 2001; Hamill et al. 2001; Whitaker and Hamill, 2002). Each assimilation includes 500 analysis cycles and the RMSE are calculated from all analysis steps after spinup (200 analysis cycle) and averaged for all the grids. 500 sets of experiments are performed and no inflation was used in the experiments.

2.7 Appendix 2C: QG-slab model

A barotropic atmosphere model (Zhang et al. 2004) based on the equation of conservation of potential vorticity is coupled with a slab ocean

$$\frac{\partial}{\partial t} \left(\nabla^2 - \frac{f_0^2}{gH_0} \right) \psi + J(\psi, \nabla^2 \psi) + J(\psi, h') + \beta \frac{\partial \psi}{\partial x} = k \left(\frac{T}{c} - \psi \right) \quad (A2.5)$$

$$H_m \frac{\partial T}{\partial t} = \lambda T + \alpha (c\psi - T) + Q_w(y) \quad (A2.6)$$

where $\psi = \frac{g}{f_0} h$ is the geostrophic streamfunction for atmosphere and T is the SST for ocean. The

$h' = \frac{f_0}{H_0} h_{topo}$ is the effect of topography for atmosphere. The mixture layer depth H_m and the

solar forcing Q_w are symmetric to the equator. H_m linearly increases with latitude from 50m to 800m and Q_w linearly decreases with latitude from 360 W/m^2 to 310 W/m^2 . The λ is the ocean damping scale. k and α are the coupling coefficient for the atmosphere and ocean. The c is a fake relationship between atmospheric barotropic streamfunction and SST. To match the ocean damping scale of ~ 3 months and an atmospheric damping scale of ~ 3 days, (λ, k, α, c) are set equal to $(1.0, 5.56 \times 10^{-5}, 2.778 \times 10^{-5}, 10^{-7})$.

For assimilation, the state variable are used the 54 (latitude) * 64 (longitude) Gaussian grid points. The atmosphere uses the leap-frog step with a spectral method in which a rhomboidal 21 truncation is applied for the transformation between spectral coefficients and grid values. And a Robert-Asselin time filter is applied to remove the computational modes (Robert 1969; Asselin 1972). Ocean is integrated using a 4-th order Runge-Kutta method. The model couples every time step with a time resolution of 30 minutes.

The experiment design is similar to Appendix 2A but all the simulations use an ensemble size 80. Because all the ocean variance comes from atmosphere in the model, only atmosphere data assimilation is applied in our results. Actually the SST error ($\sim 10^{-3}$) is negligible after the spinup of atmosphere data assimilation (300 analysis cycles). The observations cover all the atmosphere grids with a frequency 12 hours (24 time steps) and the observation error standard deviation is $10^6 \text{ m}^2 \text{ s}^{-1}$ for streamfunction (ψ). The Gaspari and Cohn (1999) covariance localization of influence radius $D=1000$ km is used in all the simulation. Each assimilation includes 600 analysis cycles and the RMSE are calculated from all analysis steps after spinup (300 analysis cycles) and averaged for all the grids. 32 sets of experiments are performed and no inflation was used in the experiments.

Tables

Table 2.1 The mean analysis RMSE for sEAKF_n, EAKF and EnKF derived using the Lorenz-96 system from 500 experiments with 80-member ensembles and different influence radius D (first column). The forcing term F=8 for all the experiments

	sEAKF ₅	sEAKF ₁₀	sEAKF ₂₀	sEAKF ₄₀	EAKF	EnKF
D=5	0.710	0.673	0.672	0.683	0.724	0.711
D=7	0.689	0.641	0.638	0.648	0.699	0.706
D=9	0.699	0.633	0.625	0.638	0.695	0.762
D=11	0.728	0.635	0.621	0.631	0.695	0.897
D=13	0.779	0.649	0.622	0.629	0.699	1.085
D=15	0.845	0.667	0.626	0.628	0.705	1.386

Table 2.2 The mean analysis RMSE for sEAKF_n, EAKF and EnKF derived using the Lorenz-96 system from 500 experiments with 80-member ensembles and different forcing F (.first column).

The influence radius is D=7 for all the experiments.

	sEAKF ₅	sEAKF ₁₀	sEAKF ₂₀	sEAKF ₄₀	EAKF	EnKF
F=2	0.042	0.037	0.036	0.036	0.036	0.038
F=5	0.333	0.318	0.320	0.328	0.360	0.340
F=8	0.689	0.641	0.638	0.648	0.699	0.706
F=10	0.801	0.739	0.733	0.744	0.799	0.833

Figures

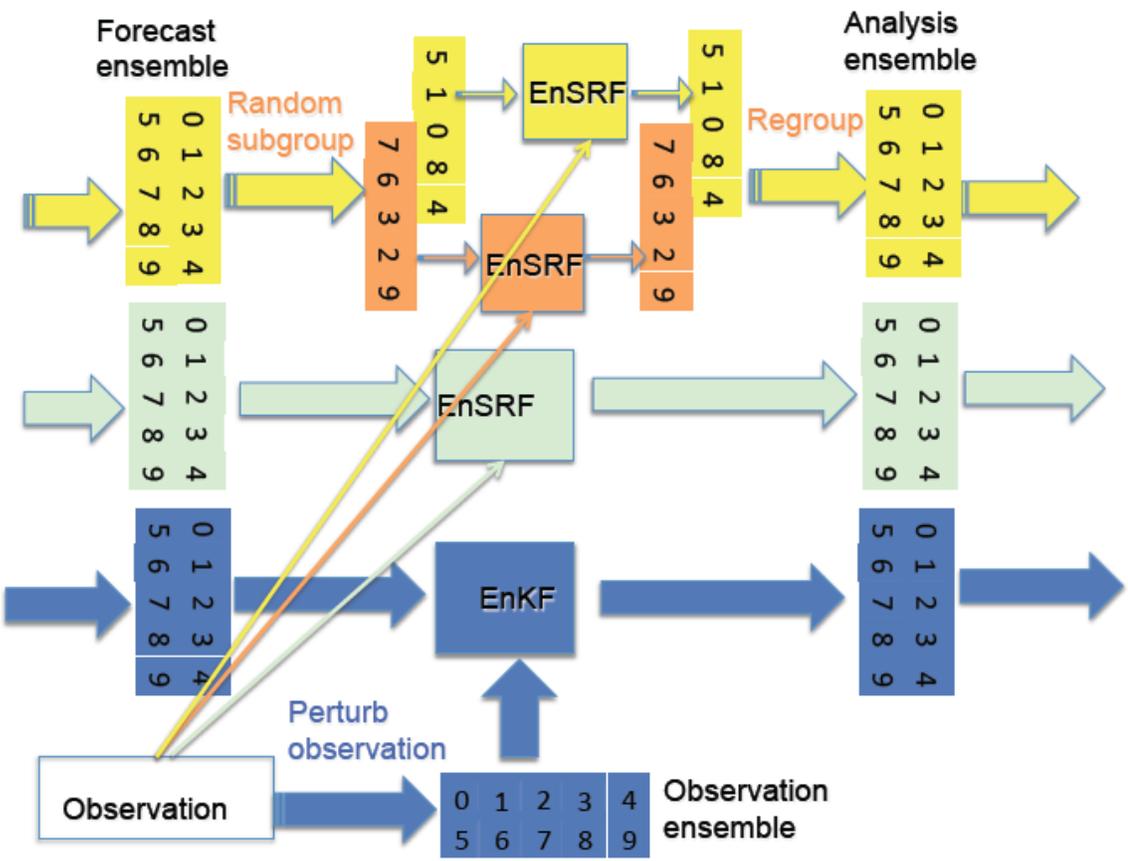


Figure 2.1 The flow charters of an analysis step for EnKF, EnSRF and sEnSRF. The blue color objects are for EnKF; the green objects are for EnSRF; and the yellow and red objects are for sEnSRF.

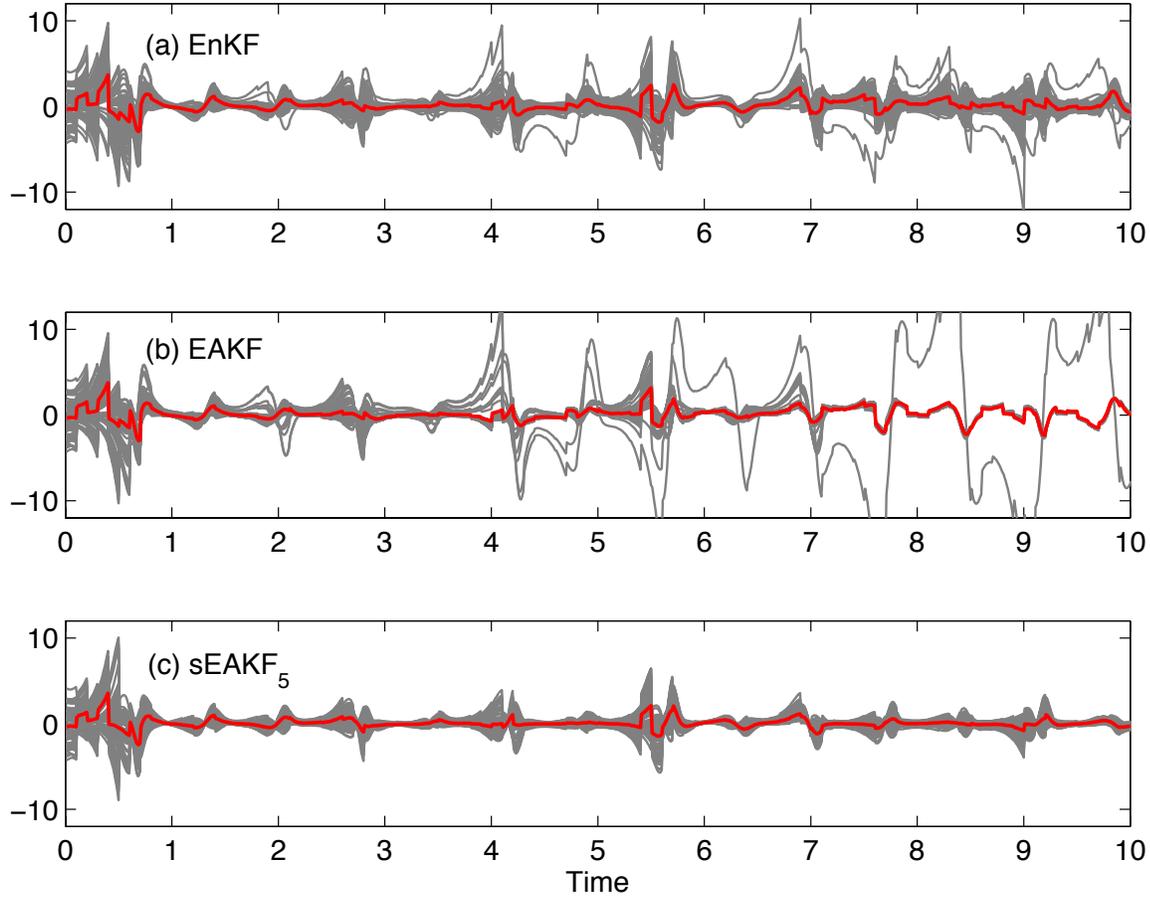


Figure 2.2 Initial error evolution of 80 ensemble members (black lines) and the ensemble means (red lines) for variable y in Lorenz-63 system with the observation time interval of 0.1 (10 analysis cycles per time unit). (a) is for EnKF simulation; (b) is for EAKF simulation and (c) is for sEAKF₅ simulation. (The figure follows figure 3 in Anderson 2010).

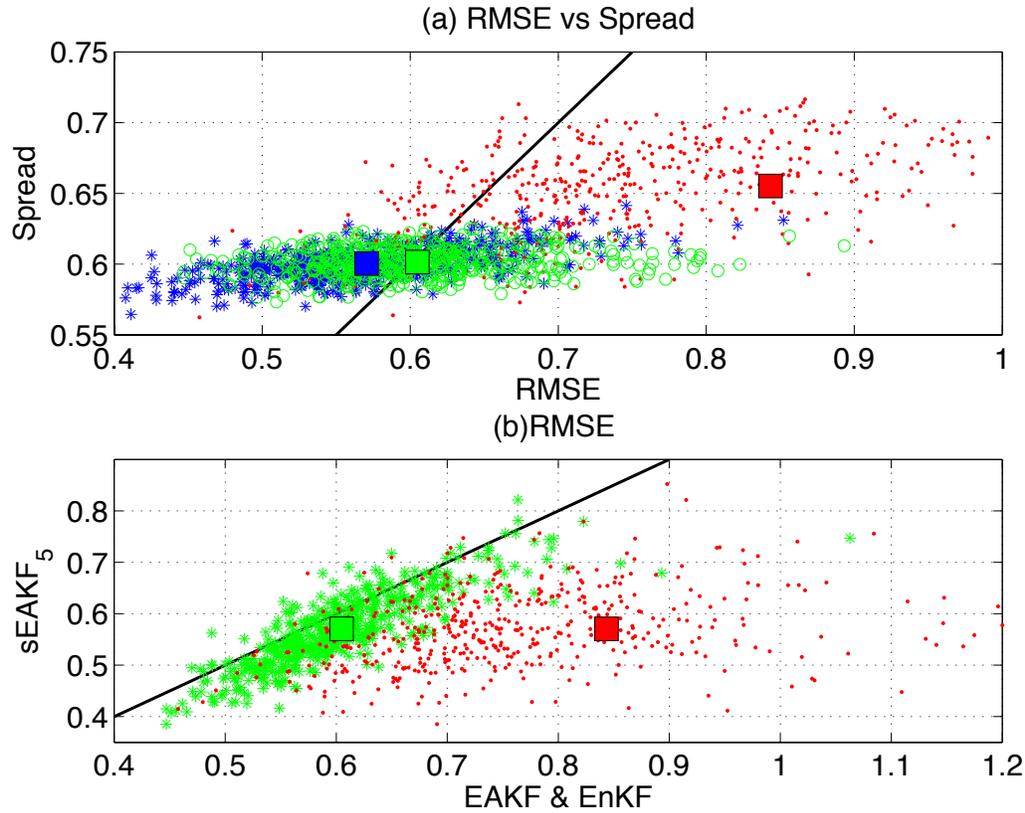


Figure 2.3 The scatter diagram for 500 sets of sEAKF₅ experiments with an 80-member ensemble and the corresponding EnKF (EAKF) simulations.

(a) The scatter diagram of analysis RMSE and ensemble spread of Lorenz-63 model for different filter schemes. The red dots are for EAKF simulation, green circles are for EnKF and blue stars are for sEAKF₅. The squares represent the average of a total of 500 experiments. A point on the black line denotes the situation when the RMSE is equal to the ensemble spread.

(b) The scatter diagram of analysis RMSE for EnKF vs sEAKF₅ (red dot) and EAKF vs sEAKF₅ (green star). The x-axis is the analysis RMSE for EAKF and EnKF simulation and the y-axis is the RMSE for sEAKF₅. The squares represent the average of total 500 sets of experiments. A point on the black line denotes when the two schemes give the same RMSE.

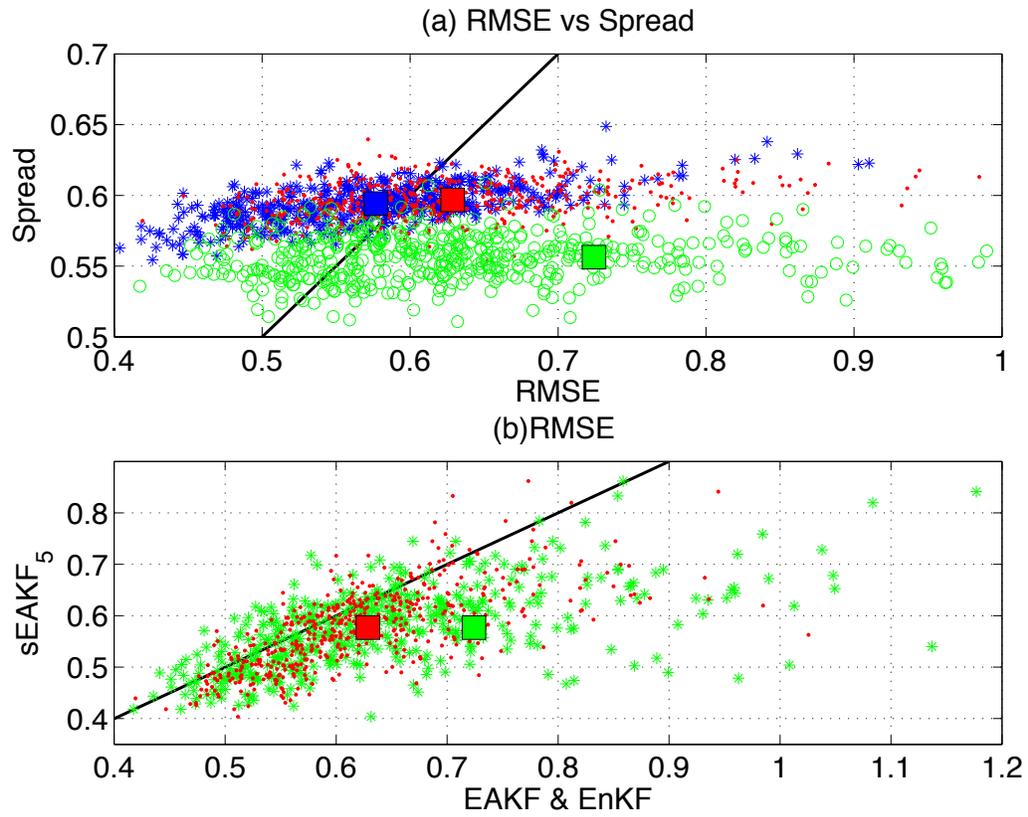


Figure 2.4 The same as fig.2.3 except for sEAKF₅ experiments with a 20-member ensemble and the corresponding EnKF (EAKF) simulations.

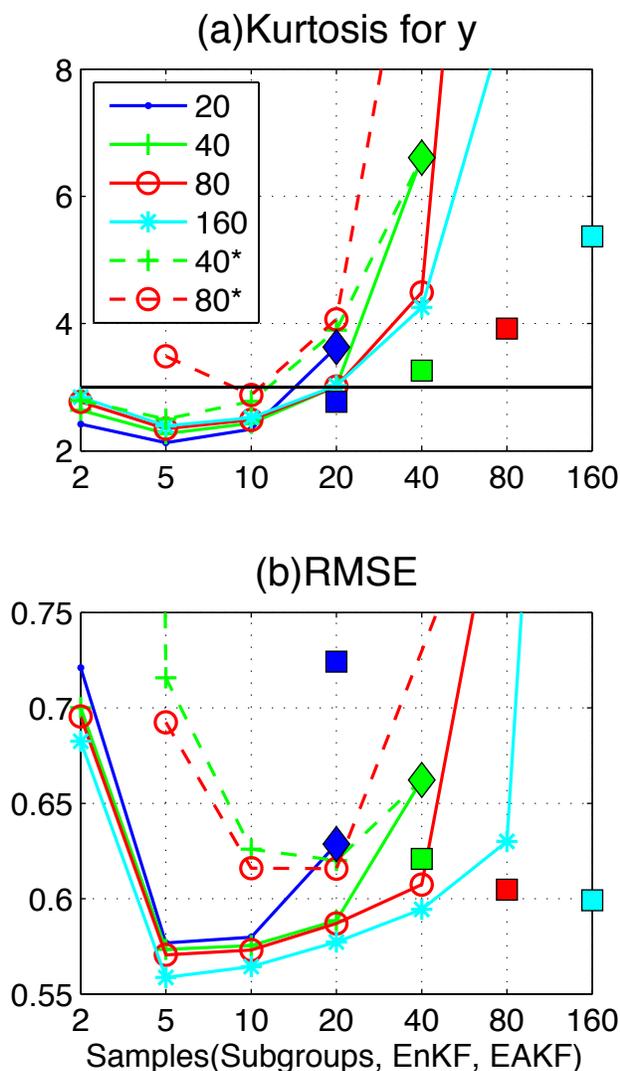


Figure 2.5 The ensemble kurtosis (a) and analysis RMSE (b) for different ensemble size averaged from 500 sets of experiments.

The x-axis represents the sample sizes for subgroups of sEAKF_n (EnKF, EAKF). The solid lines are for sEAKF_n assimilation except the black line on panel (a), which indicates the kurtosis of 3 for a Gaussian white noise distribution. The dash lines are for the simulations with subgrouping scheme but no randomization.

The blue dot lines are for 20-member ensembles; green plus lines are for 40-member ensembles; red circle lines are for 80-member ensembles; and cyan star lines are for 160-member ensembles. The squares represent the results from EnKF simulations and the diamonds represent the average results from EAKF simulations. The kurtosis (RMSE) from EAKF for an ensemble size of 80 and 160 are 19.0 (0.84) and 67.0 (1.33) that are too large to be shown on the plots.

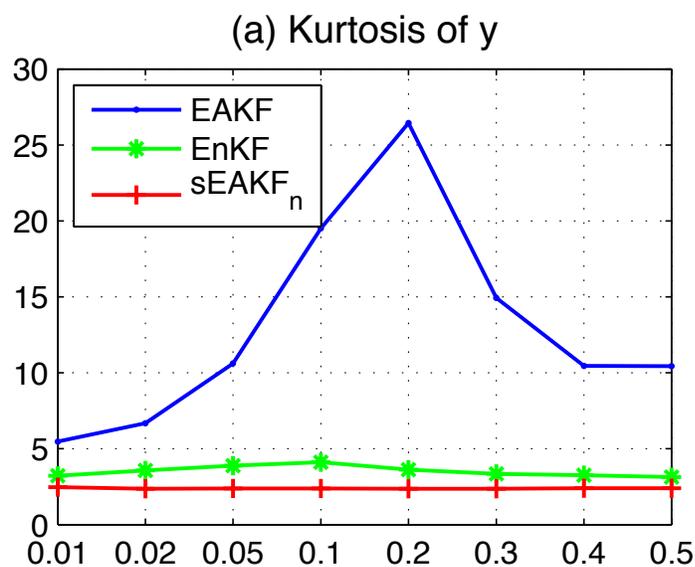
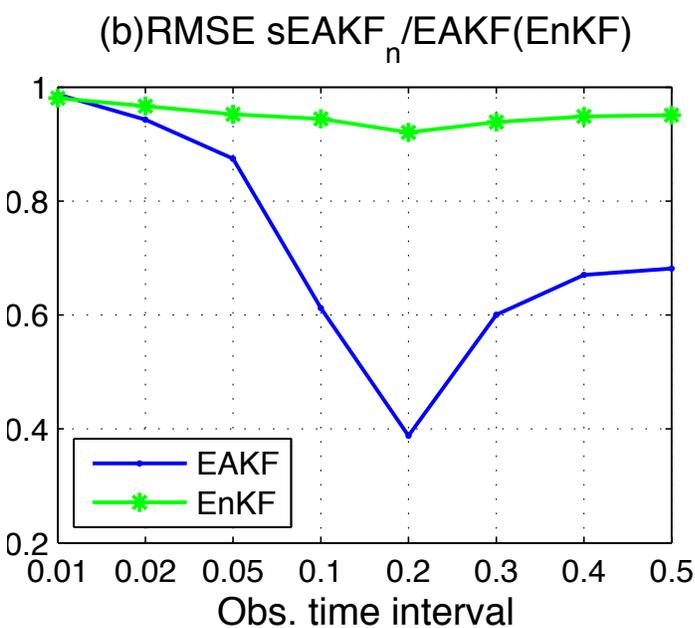


Figure 2.6 (a) The ensemble kurtosis of y for different assimilation schemes. The blue dot line is for EAKF; the green star line is for EnKF and the red circle line is for sEAKF_n with best sub-ensemble size.



(b) The RMSE ratio between sEAKF_n with optimal sub-ensemble size and corresponding EAKF (blue dot line) or EnKF (green star line).

The results are averaged from 500 sets of experiments with full ensemble size of 80 and different observation time interval. The x-axis represents the observation time intervals for the experiments.

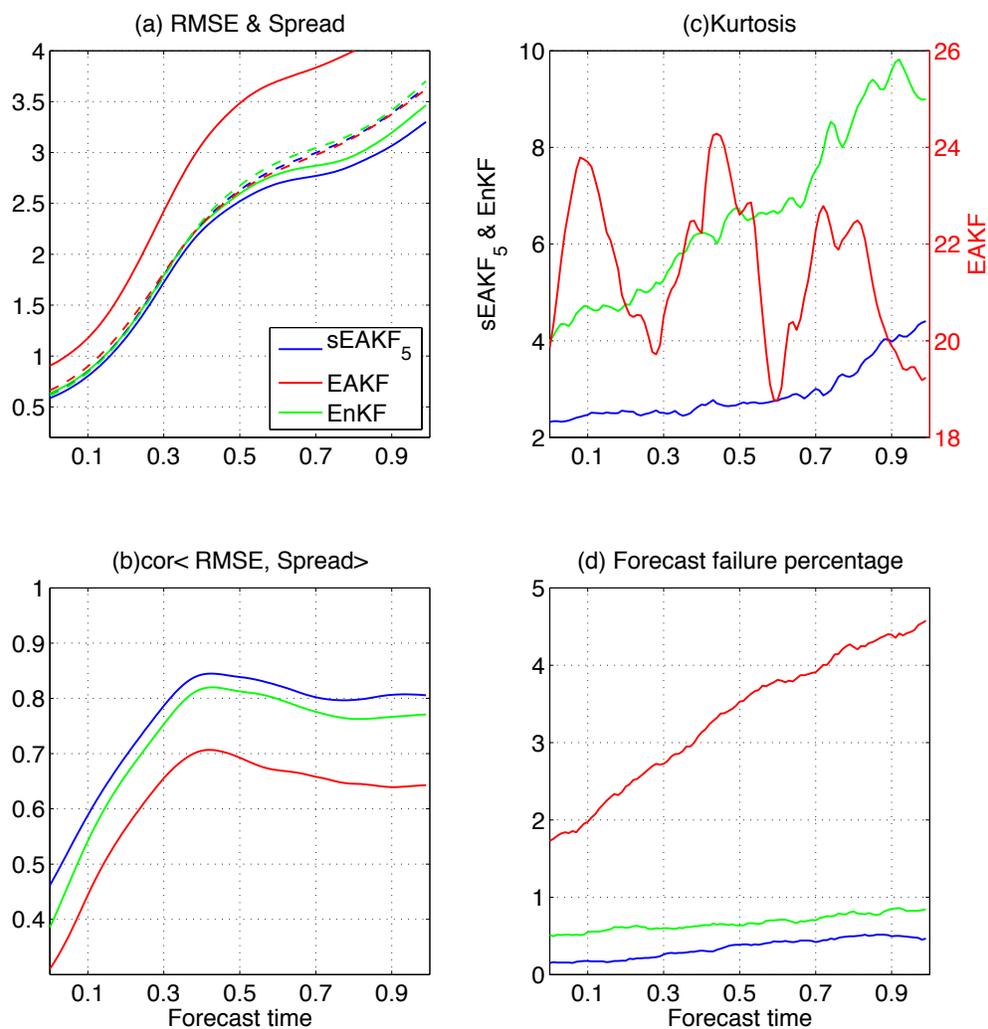


Figure 2.7 the ensemble forecasts by using initial conditions of analyses from different data assimilation schemes. The blue lines are for sEAKF₅ forecasts, the green lines are for EnKF forecasts and the red lines are for EAKF forecasts.

(a) the average forecast RMSE (solid lines) and ensemble spread (dash lines); (b) the skill-spread correlation, calculated as the correlation between the forecast RMSE and the ensemble spread for the total 20000 forecasts; (c) The average forecast ensemble kurtosis; (d) the probabilities of forecast failure in percentage, judged by RMSE greater than three times the ensemble spread.

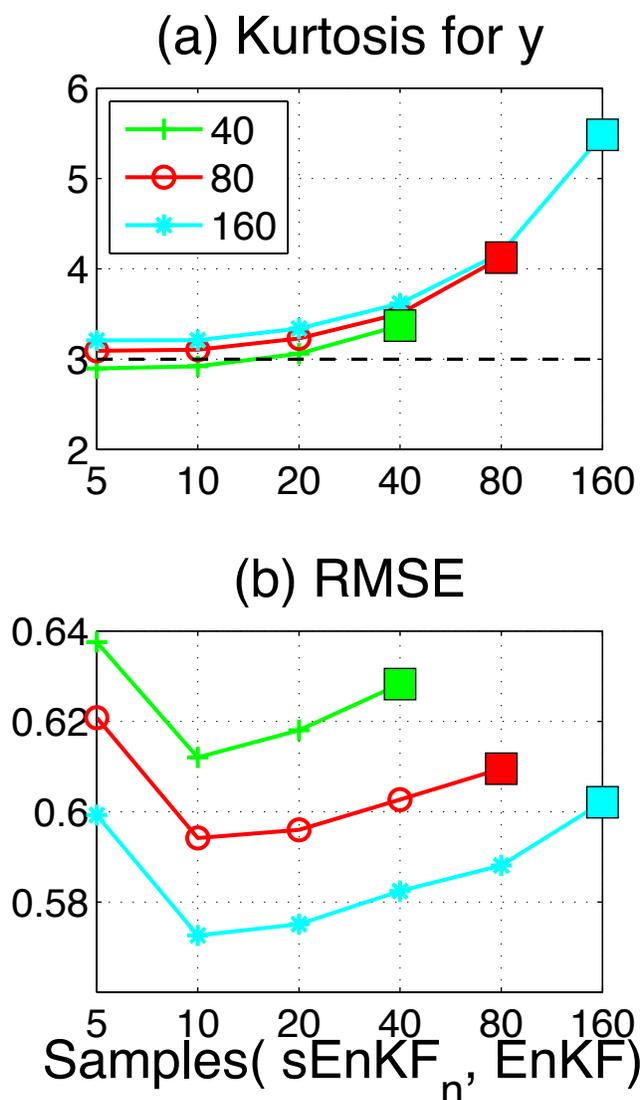


Figure 2.8 The ensemble kurtosis (a) and analysis RMSE (b) of $s\text{EnKF}_n$ and EnKF for different ensemble size averaged from 500 sets of experiments.

The x-axis represents the sample sizes for subgroups of $s\text{EnKF}_n$ and EnKF. The green plus lines are for 40-member ensembles; red circle lines are for 80-member ensembles; and cyan star lines are for 160-member ensembles. The squares represent the results from EnKF simulations. The black dash line on panel (a) indicates the kurtosis of 3 for a Gaussian white noise distribution.

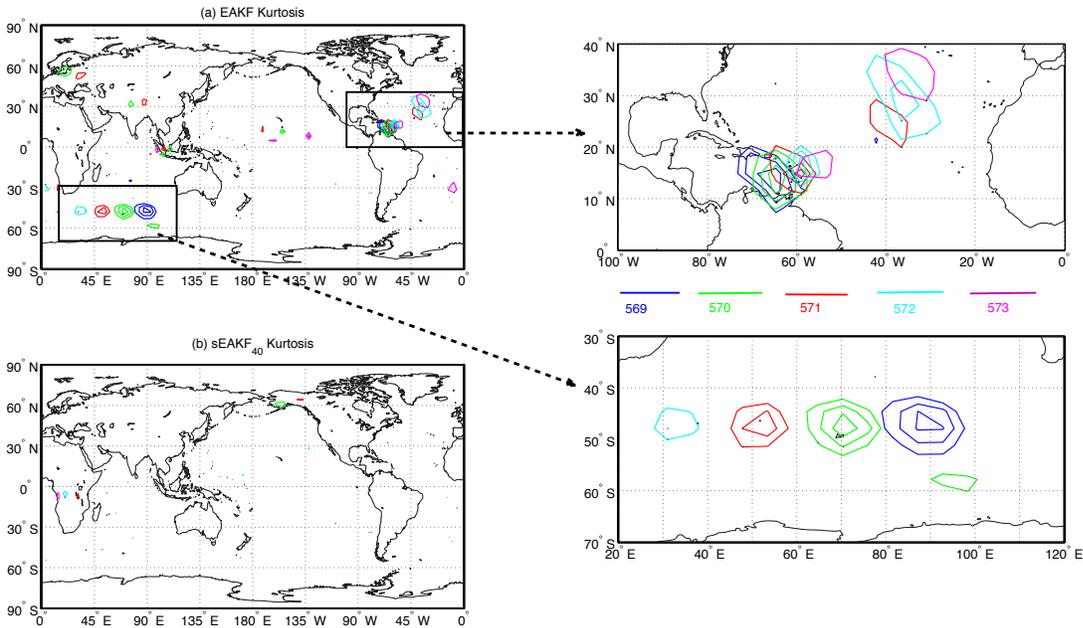


Figure 2.9 The ensemble kurtosis of ψ from one set of experiments (EAKF and EAKF₄₀), which share the identical initial condition and observation.

(a) the ensemble kurtosis of ψ with EAKF for the analysis cycles of 569-572;

(b) the ensemble kurtosis of ψ with EAKF₄₀ for the analysis cycles of 569-572;

The contour interval is 1.0 with the minimum value of 5.0. The blue contours are for analysis cycle of 569; the green contours are for the analysis cycle of 570; the red contours are for analysis cycle of 571; the green contours are for the analysis cycle of 572; the green contours are for the analysis cycle of 572.

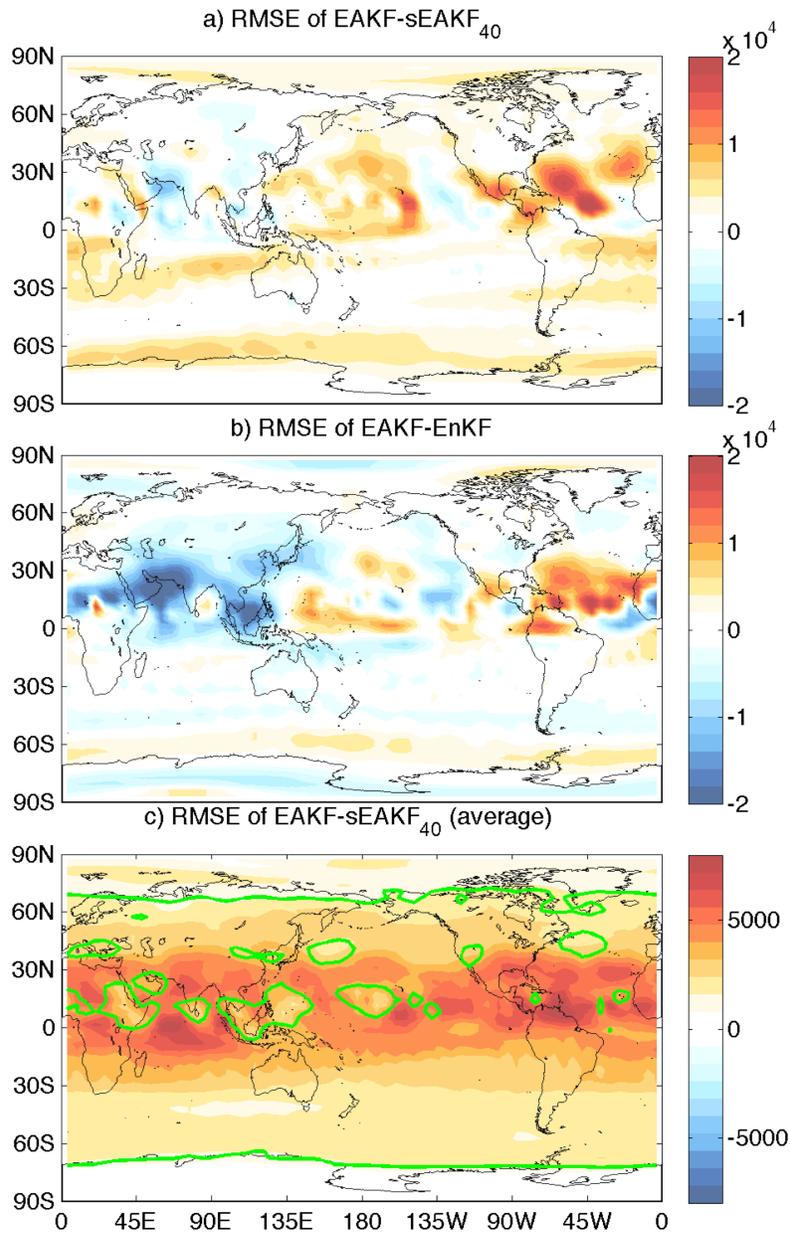


Figure 2.10 the analysis RMSE difference of ψ between different assimilation schemes.

(a) The analysis RMSE difference between EAKF and sEAKF₄₀ from one set of experiments, which share the identical initial condition and observation.

(b) The analysis RMSE between EAKF and EnKF from same set of experiment of (a).

(c) The analysis RMSE difference between EAKF and sEAKF₄₀ averaged over 32 sets of experiments. The green solid contours indicate 95% confidence level.

Chapter 3 Ensemble-based Parameter Estimation in a Coupled General Circulation Model

In this chapter, we present the first study of successful ensemble-based parameter estimation in a CGCM using an idealized observation network, demonstrating the feasibility of parameter estimation in a CGCM.

Abstract

Parameter estimation provides potentially a powerful approach to reduce model bias for complex climate models. Here, in the twin experiment framework, we perform the first parameter estimation in a fully-coupled ocean-atmosphere general circulation model using an ensemble coupled data assimilation system facilitated with parameter estimation. We first perform single parameter estimation and then multiple-parameter estimation. In the case of the single parameter estimation, the error of the parameter (solar penetration depth, SPD) is reduced by over 90% after ~40 years of assimilation of the conventional observations of monthly sea surface temperature (SST) and salinity (SSS). The results of multiple-parameter estimation are less reliable than the single-parameter estimation when only the monthly SST and SSS are assimilated. Assimilating additional observations of atmospheric data of temperature and wind improves the reliability of multiple-parameter estimation. The errors of the parameters are reduced by 90% in ~8 years of assimilation. Finally, the improved parameters also improve the model climatology. With the optimized parameters, the bias of the climatology of SST is reduced by ~90%. Overall, our study suggests the feasibility of the ensemble based parameter estimation in a fully coupled general circulation model.

3.1 Introduction

The advent of the ensemble based data assimilation scheme (EnKF) for parameter estimation (Anderson, 2001) provides a practical means for an automatic optimization of the model parameters and a reduction of model bias in a complex model. Previous studies of parameter estimation using EnKF have led to encouraging results (Annan and Hargreaves, 2004; Hacker and Snyder, 2005; Annan et al., 2005 a & b; Ridgwall et al., 2007; Hacker and Snyder, 2005; Aksoy et al., 2006 a & b; Tong and Xue 2008 a & b; Nielsen-Gammon, 2010; Hu et al., 2010; Zhang et al., 2012; Zhang, 2011 a & b; Wu et al., 2012 a & b;). Here, we will investigate parameter estimation in a CGCM using an ensemble coupled data assimilation (ECDA) scheme of DAEPC (Zhang et al., 2012) in a twin experiment framework where the parameter errors are the only source of model bias. We will show successful estimations in both cases of single parameter and multiple parameters after the parameters are carefully selected. The chapter is organized as follows. Section 3.2 briefly describes the assimilation scheme and the CGCM we used in this chapter. Sections 3.3 and 3.4 show the results of parameter estimation for single and multiple parameters, respectively. A summary is given in section 3.5.

3.2 Model and Methodology

3.2.1 The Fast Ocean Atmosphere Model (FOAM)

The model used in this study is the Fast Ocean-Atmosphere Model (FOAM), which is a CGCM with a fully parallel implementation (Jacob, 1997). The atmospheric component is a R15 spectral model with an equivalent resolution of 7.5° longitude, 4° latitude and 18 layers. The ocean component is a z-coordinate model similar to the GFDL MOM1.0 with a resolution of 2.8° longitude, 1.4° latitude and 24 layers. A simple thermodynamic sea ice model is incorporated.

Without flux adjustment, the fully coupled model has been run for over 6000 model years with no apparent drift in tropical climate (Liu et al., 2007a). In spite of its low resolution, FOAM has a reasonable tropical climatology (Liu et al., 2003), ENSO variability (Liu et al., 2000), and Pacific decadal variability (Wu et al., 2003; Liu et al., 2007b), largely comparable with current CGCMs.

3.2.2 Coupled data assimilation with enhance parameter correction (DAEPC)

In data assimilation, parameters can be estimated by augmenting state variables with model parameters (e.g. Banks, 1992a & b; Anderson, 2001). Here we will use the DAEPC (Zhang et al., 2012). The DAEPC chooses one particular EnKF scheme, the Ensemble adjustment filter (EAKF, Anderson, 2001, 2003), to estimate state variable and parameter simultaneously in the coupled system. One key factor for successful parameter estimation is to extract signal-dominant state-parameter covariance. The covariance is calculated by using the parameter uncertainty and the forecast uncertainty in the observation space. The signal is the model response to the parameter uncertainty. The noise is introduced by the limited ensemble size and is proportional to the total forecast uncertainty. Before the parameter estimation is activated, the DAEPC performs a “spin-up” process for the state estimation to reach a “quasi-equilibrium” state such that the uncertainty of the model state is sufficiently constrained by observations (see Zhang et al. (2012) for details).

3.2.3 The Adaptive Spatial Average scheme (ASA)

The parameters in this study are assumed to be globally uniform. If a globally uniform parameter is treated as a single-value parameter, there will be a large number of observations

available for updating. This will lead to the accumulation of all of the sampling errors, therefore contaminating the estimation (Aksoy et al., 2006a). To address this issue, Aksoy et al (2006a) used an updating method that transforms a globally uniform parameter into a two-dimensional field and updates the field spatially using localization. They also used a spatial average method (SA) to retain global uniformity of the estimated parameter after spatial updating. For a complex system such as a CGCM, the sensitivity and response of a model variable to a model parameter may vary spatially and temporally. Recently, the SA was further refined to an adaptive spatial average scheme (ASA) by Liu et al. (2014), which increases the convergence rate of parameter estimation in a CGCM.

Briefly, the ASA uses the ensemble spread as the criterion for selecting “good” values from the spatially varying posterior parameter field, and those “good” values are then averaged to give the final analysis of the globally uniform posterior parameter. A posterior value is called “good” if its ensemble spread significantly decreases substantially from that in the prior. The ASA calculates the uncertainty ratios between the posterior and prior. If the ratio is below a threshold, the ASA defines the posterior value as a “good” posterior value. The speed of the decrease of the parameter uncertainty depends greatly on the magnitude of the signal. Initially, the ASA can use a small value as the threshold because the initial parameter uncertainty is large and the response magnitude (signal) is large. The threshold will be increased during the simulation with the decrease of parameter uncertainty. The initial threshold in our experiments is 0.68. If the total number of “good” posterior values is less than 400, the threshold increases by 0.1 until it reaches 0.98. The ASA is applied every few EnKF analysis cycles to obtain a sufficient number of “good” parameter posterior values. The ASA is applied every 6 analysis

cycles in our parameter estimation experiments below. More detailed description of ASA will be reported in chapter 4.

3.2.4 Observations and ensemble configuration

An ensemble size of 30 is used in our experiments. A 30-year simulation from the control truth run is used for the initialization of the ensemble with the model state valid at January 1st of each year. The observations in ocean are monthly sea surface temperature (SST) and salinity (SSS), which cover the global ocean basin. The observations are generated by adding Gaussian white noise on the corresponding “truth” states at each grid point with the observational error scales (standard deviation) of 1 °K and 1 psu (practical salinity units) for SST and SSS, respectively. The ocean surface observations are used to update the upper 8 layers of ocean temperature (T) and salinity (S) (0~235m). Here the cross-covariances between SST and SSS are used to update each other. The Gaspari and Cohn (1999) covariance localization is applied with the influence radius of 3 grid points horizontally for the state variables. The observations in the atmosphere include winds and temperatures (U, V, T) for all the atmospheric grids (3D) with the error scales of 1.0 m/s, 1.0 m/s and 1.0 °K, respectively and a time interval of 12 hours. This gridded reanalysis format setting of atmospheric observations was applied in the Geophysical Fluid Dynamics Laboratory’s ECDA system (Zhang et al, 2007). An observation is used only to update the state variables (U, V, T) at its own location.

We first perform single parameter estimation and then multiple-parameter estimation. In single parameter estimation experiment, we only assimilate the oceanic observations and the biased parameter converges to the “truth” values after the assimilation. We have tested two different experimental setting for multiple parameter estimation. One experiment (EXP-M1)

only uses the oceanic observations as the single parameter estimation experiment. The other experiment (EXP-M2) assimilates both oceanic observations and atmospheric observations.

In parameter estimation, an observation can constrain a parameter directly and indirectly. An observation constrains a parameter directly by updating the parameter using the state-parameter covariance between the forecast of the observational variable and the parameter. An observation can constrain a parameter indirectly by constraining the state variables and thus improving the analysis and forecast of state variables and the state-parameter covariance. The state-parameter covariance is the key for parameter estimation and is expected to be signal-dominant. The signal, generated by the parameter uncertainty, is only part of the model total forecast uncertainty. The noise, introduced by the limited ensemble size, is proportional to the model total forecast uncertainty. The weights of the signals on the total forecast uncertainty, to some extent, indicate the signal/noise ratio of state-parameter covariance(s). The weights, which are different for different state variables, can be quantified from model forward sensitive experiments (Aksoy et al., 2006a; Tong and Xue, 2008a; Nielsen-Gammon et al, 2010). One can choose the state variable with the biggest weight to directly update a parameter to enhance the signal/noise ratio of state-parameter covariance.

Here, the SST is the chosen variable to directly update parameters for both the single parameter estimation and multiple parameter estimation. Other observations constrain the parameters indirectly through constrained the model state, which improves the forecast of SST and the state-parameter covariance. The parameter updating is activated 2 years after a “spin-up” period in which only the state variables are updated by the observations. As for the updating of state variables, the direct updating of parameter also uses covariance localization (Gaspari and Cohn, 1999) with the influence radius of 3 grid points horizontally. A conditional covariance inflation

technique (CCI) as in Aksoy et al. (2006b) is also employed here on parameter ensembles after each ASA step to avoid filter divergence for parameter estimation. The CCI inflates the parameter ensemble spread back to a predefined minimum value when necessary. The predefined minimum value is also the final uncertainty target for the estimated parameter.

3.3 Single parameter estimation

We first use the solar penetration depth (SPD) as the parameter for estimation. Solar attenuation in the ocean is a function of the amount of biomass in the upper layers of the ocean (Smith and Baker 1978, Ohlmann et al. 2000). Following Murtugudde et al. (2002), the downward solar radiation $I(z)$, at a depth of z in FOAM is calculated as

$$I(z) = I(0)re^{(-\frac{z}{h})} \quad (3.1)$$

where $I(0)$ is the total incident solar radiation at the sea surface and $r = 0.47$ (Frouin et al. 1989) represents the fraction of total solar radiation in the photosynthetically available radiation band (wavelengths from 380 to 700nm). The remaining fraction of solar radiance is fully absorbed in the top model layer of 20 meters. The h is the SPD, which will be estimated in our experiments. In the real world, the SPD can be treated as a state variable, too, because it can be calibrated using the remote sensing observation of ocean color. Here, however, it is treated as a globally uniform model parameter in FOAM that will be estimated using DAEPc.

Previous studies suggest the SPD is a parameter that has a significant impact on the surface climate (Schneider and Zhu, 1998; Nakamoto et al., 2001; Murtugudde et al., 2002; Ballabrera-Poy et al., 2007; Anderson et al., 2007). This impact can also be seen in FOAM in the difference of the climatology of SST between two simulations with different SPDs (the one with a 20-m SPD minus the one with a 17-m SPD) (Fig.3.1). A larger SPD induces significant

surface warming over the tropical Pacific, consistent with previous studies (Murtugudde et al., 2002; Ballabrera-Poy et al., 2007; Anderson et al., 2007; Hokanson, 2006). Physically, a deeper SPD allows more solar radiation to penetrate below the surface layer, leaving less shortwave radiation to heat the surface layer. This direct effect tends to generate surface cooling, opposing the surface warming in the tropical Pacific. Instead, the surface warming in the tropical Pacific is caused by an indirect effect of solar penetration, which involves momentum redistribution in the oceanic mixed layer (Murtugudde et al., 2002). Fig.1 shows that deeper SPD also leads to significant surface cooling in subtropical oceans and significant warming in the Southern Ocean at high latitudes. The indirect and direct effects discussed above combine to contribute to the locations of warming and cooling. Overall, this sensitivity experiment suggests that the model climate will vary with the SPD parameter.

The DAEP combined with ASA leads to a successful estimation of the SPD with the first guess of SPD of 20 m with an uncertainty of 3m (standard deviation) and the ‘truth’ SPD of 17 m (Fig.3.2). The SPD is not a dynamical variable. Therefore, its variance (ensemble spread) does not increase during the model integration; yet, its variance is reduced at each analysis step. As a result the ensemble spread of SPD initially decreases much faster than its root-mean-square error (RMSE) (Fig.3.2). The CCI prevents parameter variance from decreasing indefinitely by adopting a minimum parameter ensemble spread of 0.3m (1/10 of the initial standard deviation) in the first 30 years of simulation. The minimum parameter ensemble spread is decreased to 0.2m for the simulation afterwards (year 31~47), when we believe the estimated SPD has converged close to the ‘truth’ value. The error of the parameter SPD decreased from 3m to 0.2m after 47 years of assimilation.

The consistency between the ensemble spread and its represented forecast error is a very important factor for EnKF to succeed. The spatial pattern of the monthly average SST error (RMSE) shows maximums at both the high latitudes and equatorial region, and a minimum in the off-equatorial regions (Fig.3.3a). This pattern resembles that of the SST variance (figure not shown) because the regions of higher variance usually have larger forecast errors. The larger SST RMSE and variance located in the Pacific equatorial region are related to the model ENSO variability. The spatial pattern and amplitude of the ensemble spread of SST resemble closely those of the forecast RMSE of SST (Fig.3.3b), suggesting a good quality of the ensemble-based filter.

The improvement of the parameter also improves the model climate and, in turn, the forecast errors of state variables. The experiment after parameter estimation produces a better forecast of monthly SST (the 1st month) in comparison with a pure data assimilation experiment, which uses the same experiment design but with the biased SPD parameter (20 m) and no parameter correction. The spatial patterns of the RMSE of the 1st month SST forecast resembles closely that in the experiments of pure data, but the amplitude is reduced by 12% in the former relative to the latter (0.40 °K vs. 0.45 °K) for the SST RMSE averaged between 60S and 60N. In the pure data assimilation experiment the average RMSE of 500-mb geopotential height (GPH) is ~21m, which is ~2/3 of that in a model ensemble simulation that does not assimilate any observations. The addition of parameter estimation in the ocean does not further improve the quality of atmosphere analysis and forecast.

The robustness of our parameter estimation is confirmed with another experiment that uses the SA method of Aksoy et al (2006a). The SA experiment also estimates the parameter successfully, but with a slower convergence rate (see chapter 4 for details).

3.4 Multiple parameter estimation

In this section, we extend the parameter estimation from a single parameter (SPD) to three biased parameters: one is in the ocean component and the other two are related to air-sea coupling (table 3.1). The imperfect ocean parameter is still the SPD that has been discussed in the previous section. The imperfect parameters for air-sea coupling are two artificial parameters, m_d and m_q , which are the multipliers to the momentum and latent heat fluxes, respectively, between the ocean and atmosphere (calculated in the coupler component of the model). Thus, $m_d = 1.0$ and $m_q = 1.0$ recover the default setting of air-sea coupling. The specific value and the minimum ensemble spread for each imperfect parameter are shown in table 3.1.

The model climatology of SST shows significant sensitivity to the two coupling parameters, m_d and m_q (Fig.3.4). The parameter m_d directly influences the momentum flux between the ocean and the atmosphere. When m_d is increased from 1 to 1.2, the SST shows a significant warming in the subtropical oceans and cooling at higher latitudes (Fig.3.4a). The warming in the subtropics seems to be induced, partly, by the slower surface wind (in response to a larger drag coefficient) and in turn reduced evaporative cooling, while the cooling in the mid-latitude and subpolar region may be contributed by a stronger mixing of the colder water from the bottom of the mixed layer and a stronger Ekman upwelling. The parameter m_q influences the latent heat flux between the atmosphere and ocean, and therefore impacts SST directly. An increase in m_q (from 1 to 1.2) enhances latent heat flux cooling and therefore leads to a significant surface cooling over the global ocean, except for high latitudes where the latent heat flux is small (Fig.3.4b).

The results of multiple-parameter estimation are not as good as the single-parameter estimation when the same experimental setting is used, because the nonlinearity between parameters and state variables weakens the correlation between forecast error and individual parameter uncertainty. For the EXP-M1, some parameters do not converge when we only assimilate the oceanic observations (monthly SST and SSS) (Fig.3.5). Only the SPD successfully converges to the “truth” in 50 years of assimilation (Fig.3.5a). The evolution of estimation SPD has slower convergence speed and less parameter stability, compared with the estimation SPD in single parameter estimation (Fig.3.2). The error of the m_q is only reduced by ~50% in 50 years of assimilation (Fig.3.5c) and the error of the m_d shows slight increase after the assimilation (Fig.3.5b). However, the m_d and m_q converge successfully in the single parameter estimation with the same observational setting (see figure 4.5 in chapter 4). The lower estimation performance of multiple parameter estimation, compared with single parameter estimation, is consistent with previous works (Aksoy et al., 2006b; Tong and Xue, 2008b; Hu et al., 2010).

The smaller reliability of multiple-parameter estimation can, in theory, be improved by decreasing the forecast errors of model variables that are used to constrain parameters directly. The assimilation of additional atmospheric observations into the model in EXP-M2 generates more accurate analysis and forecast of SST. The smaller forecast uncertainty of SST, with the reduced sampling error, enhances the signal/noise ratio of state-parameter covariance, which, then, accelerates the convergence of the parameter estimation (Fig.3.6). The parameters in EXP-M2 almost all converge to the “truth” values after a 16-year assimilation and the convergence speed is much faster than the speed of those in EXP-M1 (Fig.3.5). The estimated m_d and m_q monotonically converge to the “truth” in 8 model years (Fig.3.6b & c). The estimated SPD

initially exhibits a small overshoot, i.e. the parameter error decreases from positive 3.0 m to negative ~ 0.5 m, and then converges back to the “truth”. The estimated SPD relaxes back to the truth in 8 model years (Fig.3.6a). Our other experiments show that this type of overshooting sometimes occurs, yet appears to have little impact on the final convergence. Similar to the single parameter estimation that has been discussed in section 3, parameter ensemble spreads initially all suffer a negative bias, compared with their RMSEs, which confirms the necessity of applying CCI on the parameter ensemble spreads.

The parameter estimation also helps to improve the analysis of the state variables. Here we use the short forecast to indicate the decrease of analysis error of state variables (because the analysis are not saved in our experiment). Fig.3.6d shows the evolution of the forecast RMSEs of monthly SST and 500-mb geopotential height (GPH) during the assimilation. During the “spin-up” period of DAECF (first two years), the forecast errors for SST decrease very rapidly and reach the quasi equilibrium in a few months, with the average RMSE reduced from $\sim 1^{\circ}\text{K}$ to $\sim 0.2^{\circ}\text{K}$. The RMSE further decreases to $\sim 0.1^{\circ}\text{K}$ when the parameter updating is activated and the uncertainties of parameters are reduced. During the period of parameter estimation (after year 2), the ensemble spread of forecast SST becomes smaller than the SST RMSE (Fig.3.6d), but the smaller ensemble spread of state variables does not seem to affect the parameter estimation. All the three biased parameters converge to the “truth” values quickly. It should be pointed out that, for simplicity, we have not used covariance inflation on state variables to enhance their ensemble spreads. It is conceivable that, with an inflation scheme, the estimation of the parameter and state variables will be further improved.

Similar to SST, the forecast error of GPH decreases dramatically with the assimilation of atmospheric observations of U, V and T (Fig.3.6d). The forecast error reaches the quasi

equilibrium during the “spin-up” period of DAECF with an average RMSE of ~ 2.0 m. The RMSE further decreases to ~ 1.0 m when the parameter updating is activated and the uncertainties of parameters are reduced. The ensemble spread of GPH is sensitive to the parameter ensemble spread. The ensemble spread is greater than its RMSE during the “spin-up” period of DAECF and becomes smaller than its RMSE when the parameter updating is activated and the parameter ensembles suffer the negative bias. The ensemble spread of GPH and its RMSE become consistent when the analysis reach equilibrium and the negative bias of parameter ensemble disappear. In addition, to decrease the computational cost, the atmospheric observations are only used to update the state variables locally in EXP-M2. It is expected that the state variables will be further improved when the observations are also used to update the nearby regions with a covariance localization scheme.

As expected, the improved parameters also improve the model climate. The bias of the SST climatology, generated by the initial parameter errors (see table 3.1), shows significant cooling with an average RMSE of ~ 0.61 °K (Fig.3.7a). The spatial pattern is very similar to Fig.3b because the effect of m_q is the strongest among the three biased parameters. The weak bias along the equatorial region is due to the warming produced by the positive biases of SPD and m_d , which counteracts the cooling generated by the biased m_q . The significant cooling of SST accompanies a cold bias in the atmosphere in FOAM, which lowers the GPH (Fig.3.8a). The GPH climatology at 500-mb shows a significant negative bias with an average RMSE of ~ 12.5 m. The spatial pattern of GPH bias (Fig. 8a) matches the spatial pattern of SST climatology bias (Fig.3.7a). When the updated parameters (table 3.1) are used, the biases in SST and GPH climatology decrease dramatically with the RMSE of SST and GPH reduced to ~ 0.05 °K and ~ 1.4 m, respectively (Fig.3.7b and Fig.3.8b). Overall, like the single parameter

experiment, the multi-parameter estimation also improves the model climate and forecast significantly.

3.5 Summary

In this study, we explored the parameter estimation in a CGCM using an ensemble-based assimilation scheme in the twin experiment framework. Here, our DAEPC successfully optimized the single imperfect parameter SPD using the conventional observations of monthly SST and SSS. The SPD error was reduced from 3-m to 0.2-m after ~40 years of assimilation. The DAEPC also performed well in the experiment of multiple parameter estimation by using the 12-hourly atmospheric winds and temperature observations and the monthly SST and SSS observations in the ocean. The three imperfect parameters all converged on the “truth” values after a 16-year assimilation.

The improved model parameter also improved the model climatology and model forecast. The RMSE of the SST climatology was reduced from ~0.6K to ~0.05K, from the model of initial biased parameters to the optimized parameters (table 3.1). The RMSE of the forecast monthly SST (1st month) was reduced by 12% with the parameter correction in the experiment of single parameter estimation. The forecast RMSE of monthly SST and GPH decreased from ~0.12 °K to ~0.07 °K and from ~2.0 m to ~ 1.0 m, respectively, by correcting biased parameters in the experiment of multiple parameter estimation.

It is important for ensemble based parameter estimation to choose the right observations to update parameters directly. Parameters are not dynamical variables; they cannot be modified by model dynamics, but only by the observations directly through the state-parameter covariance. The error of a parameter decreases when the parameter is directly updated by the

observations with signal dominated state-parameter covariance(s). The error of a parameter could increase when the parameter is directly updated by the observations with a noise dominated state-parameter covariance(s). The signal/noise ratios of state-parameter covariance(s) are different for different observational variables. To retain successful parameter estimation, we have to choose the observational variables with robust state-parameter covariance(s) to directly update the parameters. The other observations can still improve parameter estimation by improving the model forecast and enhancing the state-parameter covariance used for parameters updating. In this study, SST was chosen as the observational variable to directly update parameter, which leads to the parameter estimation success in single parameter estimation. When multiple parameters are biased in EXP-M1, the state-parameter covariance(s) of SST become less robust, which results in a failure to estimate the parameters of m_d and m_q . Decreasing the forecast error of SST can enhance the state-parameter covariance and improve the parameter estimation. Indeed, The assimilation of additional atmospheric observations into the model in EXP-M2 narrows the SST uncertainty and produces successful estimation for all three biased parameters. However, when we replaced SST as SSS to directly update parameters, the parameter estimation failed even for single parameter estimation, because the response of SSS to parameters is much weaker than that of SST.

To our knowledge, this is the first demonstration of successful ensemble-based parameter estimation in a general circulation model with fully coupled ocean-atmosphere dynamic. It demonstrates the feasibility of parameter optimization in a complex CGCM using an ensemble-based filter for parameter optimization and therefore suggests the potential of parameter optimization to reduce model bias and improve CGCMs in the future. The idealized observation network used in this study is very different than the realistic observation network. The

atmospheric observations applied in this study are using the gridded reanalysis format setting. Previous works show that, by assimilating the reanalysis data, ECDA significantly improves the forecast and analysis in Geophysical Fluid Dynamics Laboratory Climate Models (Zhang et al, 2007; Yang et al., 2013; Chang et al., 2013; Zhang et al, 2014). The ensemble-based parameter estimation in the CGCM using realistic reanalysis products as observations remains to be further studied.

3.6 Appendix: An experiment with unsuccessful parameter estimation

We have demonstrated the successful parameter estimation in this chapter with the carefully selected test parameters. However, it is not guarantee that all the biased parameters can be estimated using EnKF. Our EnKF experiment (EXP-M3) fails to estimate a mixing parameter v_s in FOAM with the same experiment design in the multiple parameter estimation EXP-M2, which assimilate both atmospheric and oceanic observation.

FOAM uses a Richardson-number based vertical mixing parameterization scheme, which is the PP mixing scheme (PP 1981) modified after Peters et al. (1988). The equation for the vertical mixing coefficient for temperature and salinity is

$$K_{ts} = \text{Min} \left(\frac{v_s}{(R_i)^8} + k_b, 0.01 \right) \quad (\text{A3.1})$$

where R_i is the local gradient Richardson number and $k_b = 1.0 * 10^{-5} m^2/s$ is the background diffusivity.

The EXP-M3 estimates the bias v_s as well as the SPD, m_q and m_d at the same time. The results of estimation show that the DAEPC with ASA method successfully optimizes the parameters of SPD, m_d and m_q , but not v_s (Fig.3.9). The estimation of v_s does not converge towards the “truth”. Rather, its value decreases monotonically from 3.05 to 2.50 in a 20-year

assimilation while its ensemble spread decreases dramatically. It therefore appears that this parameter suffers from filter divergence. We are not completely clear why the estimation of ν_s fails. One reason is, we believe, the strong nonlinear threshold feature of the vertical mixing coefficient. The response of the vertical mixing K_{ts} to the ν_s uncertainty is modulated by a maximum threshold of $0.01m^2/s$ (see eqn. (A3.1)). Once K_{ts} reaches this threshold, it will not be affected by the variation of ν_s any more. As such, it is difficult to find a clear linear relationship between the SST forecast ensemble and the ν_s ensemble. More studies are needed on this issue.

The failure of the estimation of ν_s indicates the limitation of ensemble based parameter estimation. We have to avoid the limitation by choosing the estimated parameters with less nonlinearity. In the mean time, the failure of the estimation of ν_s and the success of the estimation of the other three parameters, nevertheless, may have an important implication to more general multi-parameter estimation and, even the real world application.

Table

Table 3.1 Multiple parameter estimation experiment. (The estimated values are come from EXP-M2.)

Parameter (unit)	Initial guess value	Estimated value	Truth value	CCI threshold
SPD (m)	20.0	17.1	17.0	0.3
m_d	1.20	0.99	1.0	0.02
m_q	1.20	1.00	1.0	0.02

Figures

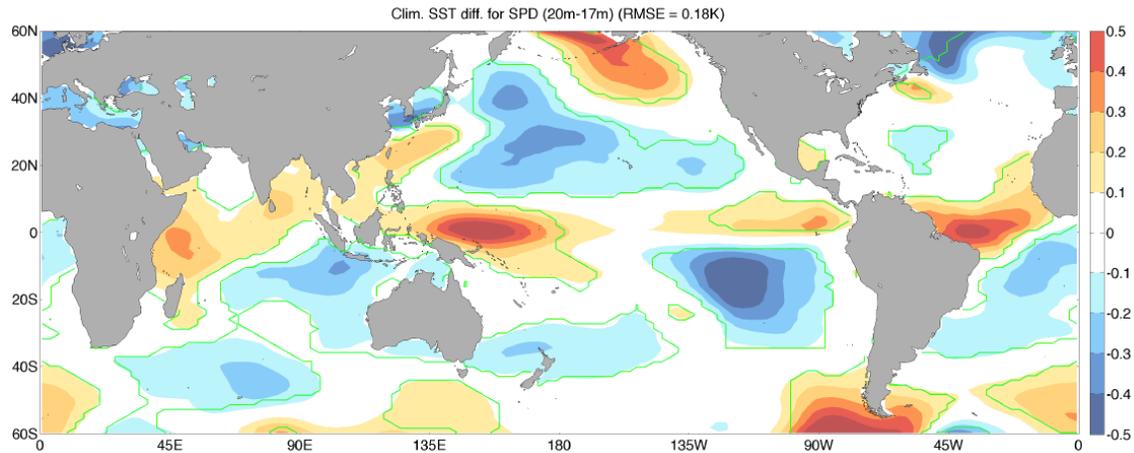


Fig. 3.1. The climatological annual mean SST difference between two simulations with the same initial conditions but different SPD (20m -17m). The climatological mean SST(s) are calculated from the 80-years average after 20 years of spin up. The green contours represent the 95% confidence level.

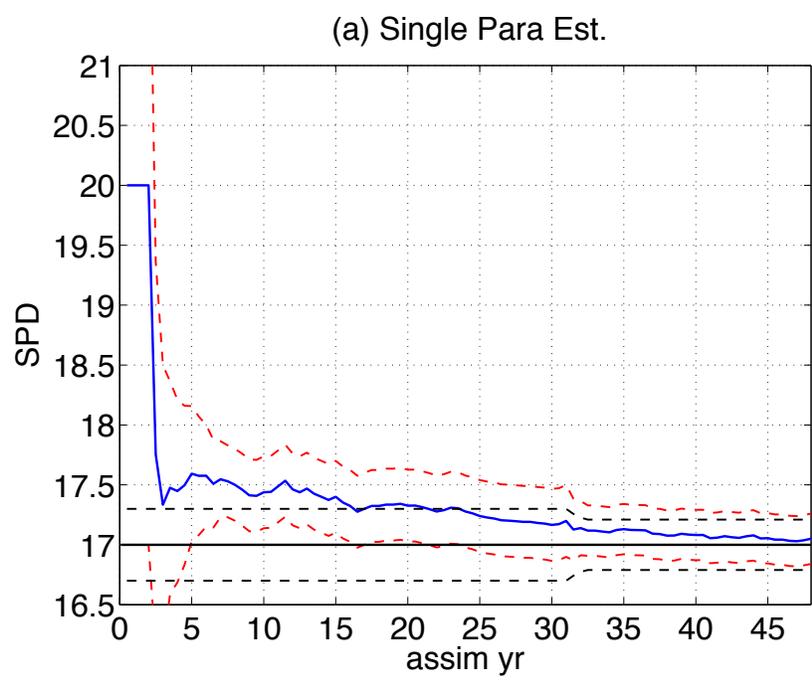


Fig. 3.2. Single parameter estimation (SPD) uses DAEPC with the adaptive spatial average (ASA) method. The ensemble size is 30. The observations are the monthly SST and SSS. The blue line is for the temporal evolution of the ensemble mean of SPD and the red dashed lines are the 1-standard deviation of its ensemble spread. The black solid line is the “truth” and the black dashed lines are the minimum parameter ensemble spreads (uncertainty goals) of CCI in the experiment.

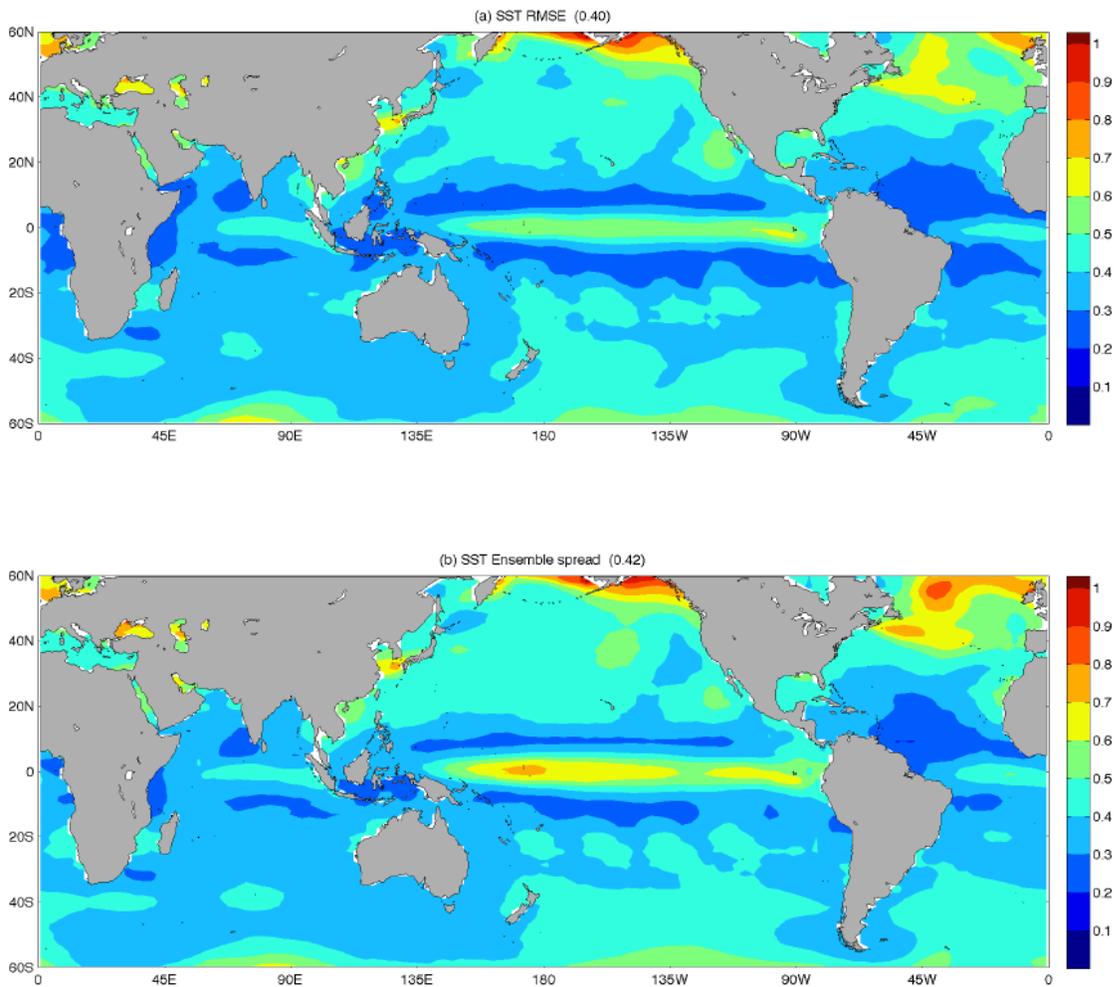


Fig. 3.3. The spatial patterns of the (a) RMSE and (b) ensemble spread of the forecast monthly SST (1st month) averaging 332 cases, which initiate from each month of the simulation years of 21-48 in the single parameter estimation experiment using DAEPC with the adaptive spatial average method.

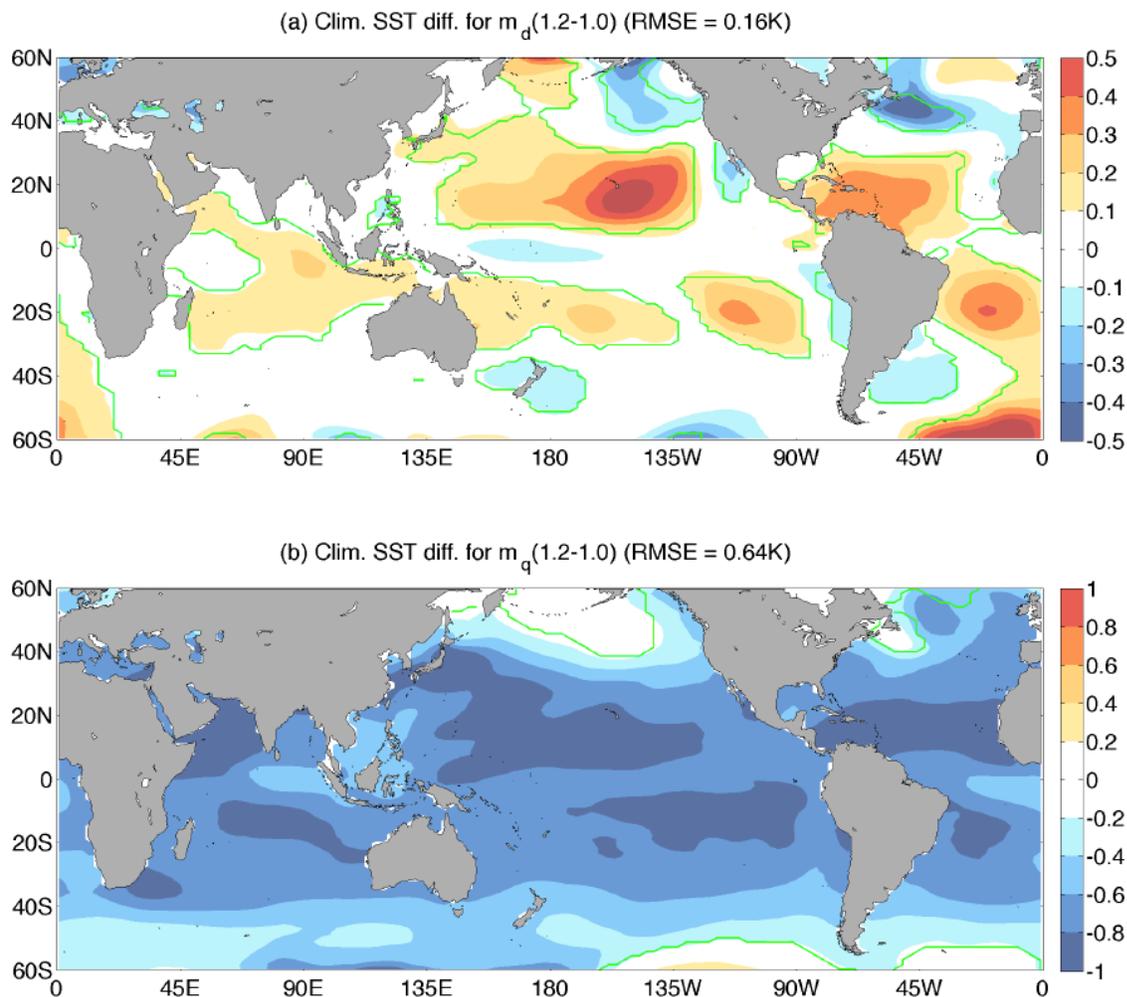


Fig. 3.4. The climatological annual mean SST difference between two simulations with the same initial conditions but different parameters: (a) the different m_d (1.2 - 1.0); (b) the different m_q (1.2 - 1.0). The climatological mean SST(s) are calculated from the 80-year average after 20 years of spin up. The green contours represent the 95% confidence level.

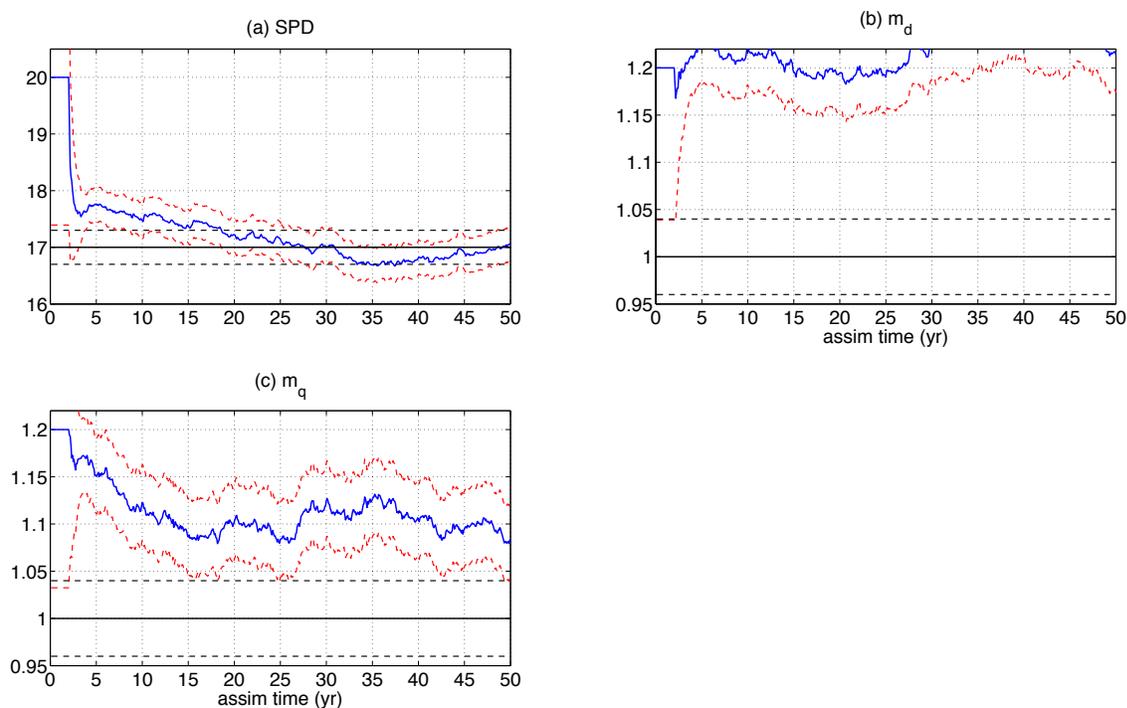


Fig. 3.5. The temporal evolution of estimated parameters for the multiple parameter estimation experiment of EXP-M1, which only assimilates the oceanic observations of monthly SST and SSS. The panel (a), (b) and (c) are for the SPD, m_d and m_q , respectively.. The solid blue lines are the parameter ensemble means and the red dashed lines are the 1-standard deviation of ensemble spreads; the black solid lines are the “truth” and the black dashed lines are the minimum parameter ensemble spreads of CCI in the experiment.

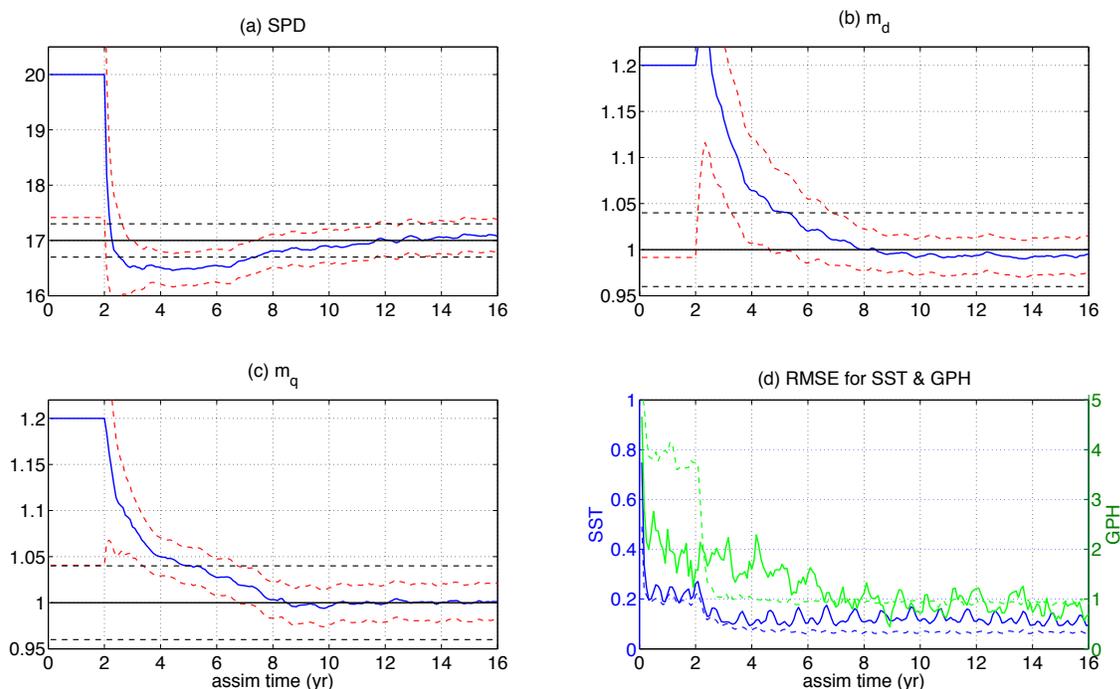


Fig. 3.6. The temporal evolution of estimated parameters and forecast RMSE of SST and GPH for the multiple parameter estimation experiment of EXP-M2, which assimilates both the oceanic observations and the atmospheric observations. The panel (a), (b) and (c) are for the SPD, m_d and m_q , respectively. The panel (d) is for the SST RMSE. The solid blue lines in (a), (b) and (c) are the parameter ensemble means and the red dashed lines are the 1-standard deviation of ensemble spreads; the black solid lines are the “truth” and the black dashed lines are the minimum parameter ensemble spreads of CCI in the experiment. The solid line and the dashed line in panel (d) are the 1 monthly forecast RMSE and ensemble spread of SST (blue lines) and GPH (green lines), respectively. The RMSE and ensemble spread are first calculated in each grid and then averaged from 60S to 60N.

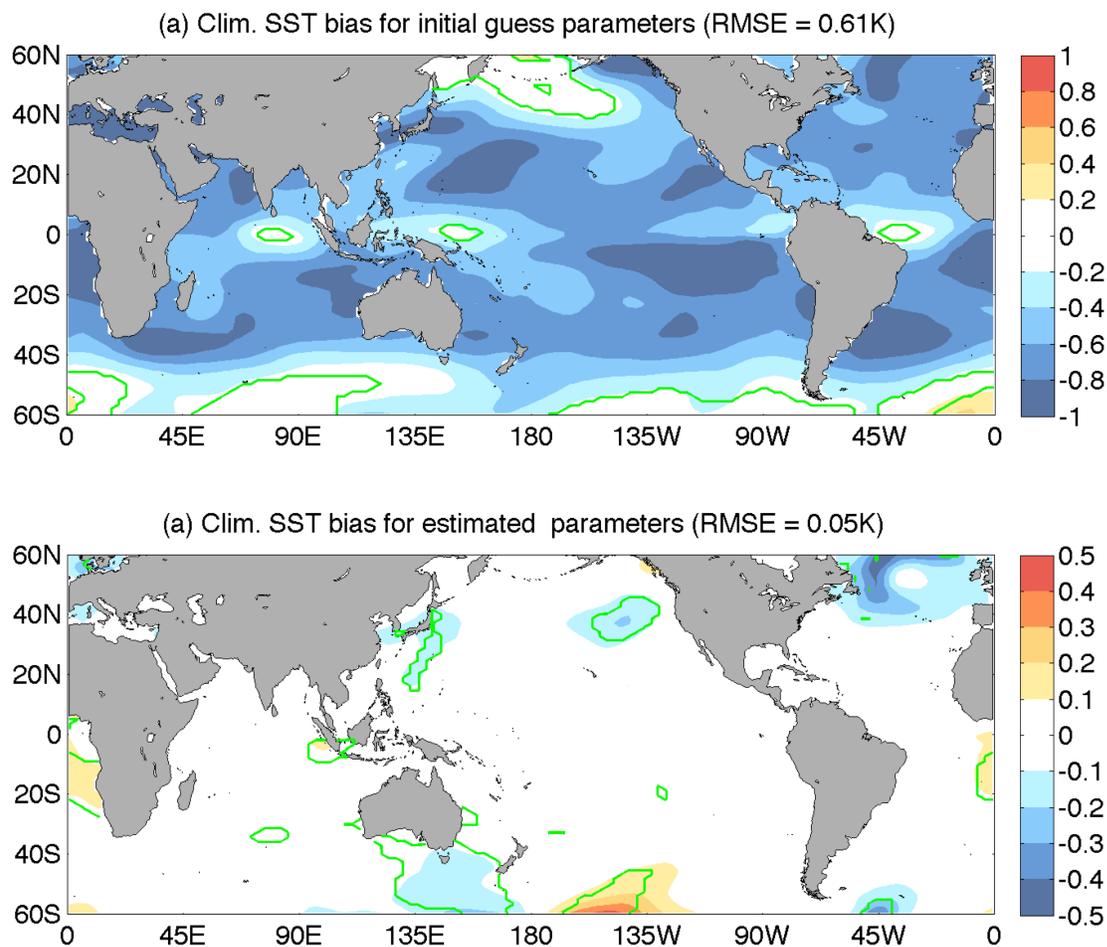


Fig. 3.7 The climatological annual mean SST difference between two simulations with the same initial conditions but different parameters: (a) the initial parameters (Table 1) minus the “truth” parameters; (b) the estimated parameters (table 1) minus the “truth” parameters. The climatological mean SST(s) are calculated from the 80-year average after 20 years of spin up. The green contours represent the 95% confidence level.

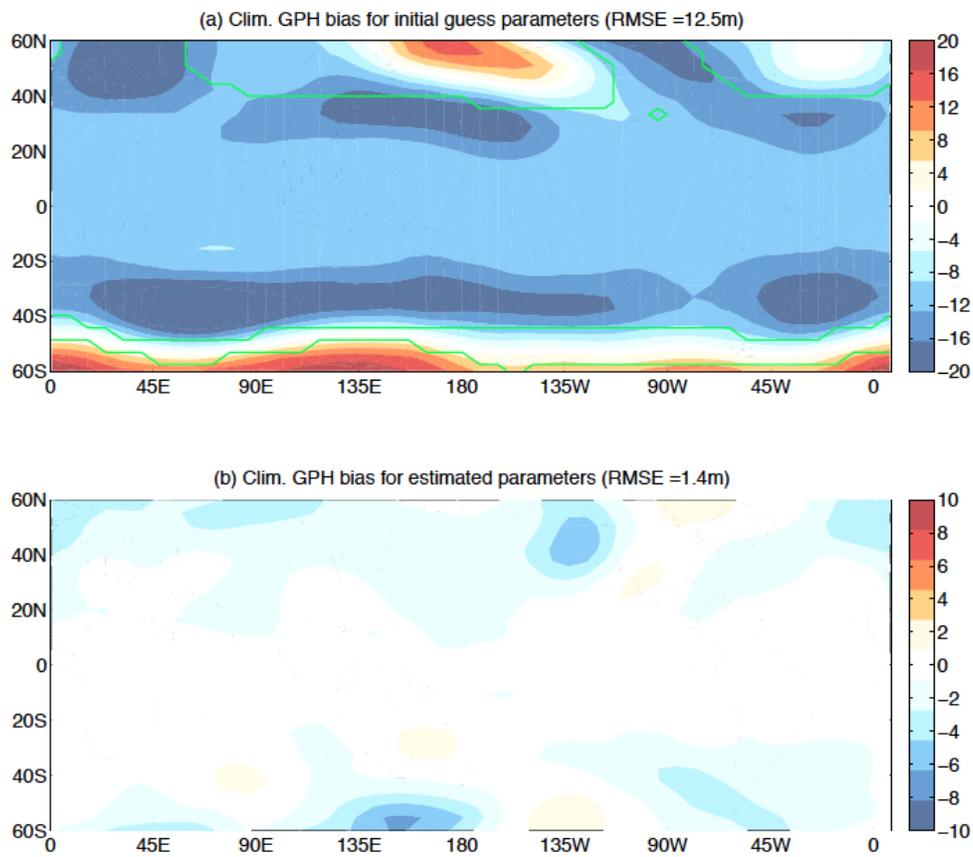


Fig. 3.8. The same as Fig. 3.7 except for GPH climatology difference at 500-mb

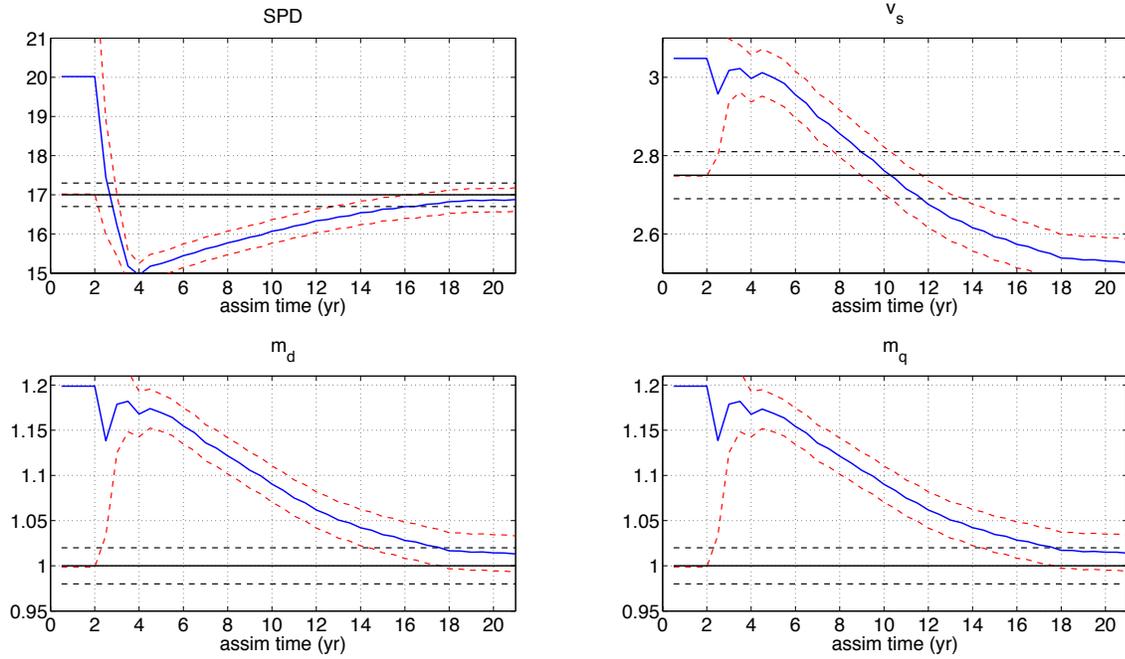


Fig. 9 The temporal evolution of estimated parameters for the multiple parameter estimation experiment of EXP-M3. The solid blue lines are the temporal evolution of parameter ensemble mean and the red dash lines are the 1-standard deviation of its ensemble spread. The black solid line is the “truth” and the black dash lines are the minimum parameter ensemble spreads (uncertainty goals) for the experiment.

The panel (a) is for parameter SPD, panel (b) is for parameter v_s , panel (c) is parameter m_d and panel (d) is for parameter m_q .

Chapter 4 Ensemble-based Parameter Estimation in a Coupled GCM Using the Adaptive Spatial Average Method

The efficiency is very important for parameter estimation in a CGCM because we only have limited observation and computation resource available. In chapter3 we applied a new cost and the limitation of observation. We have proposed an adaptive spatial average method (ASA) to improve the efficiency of parameter estimation. The ASA has been applied in our parameter estimation experiments in chapter. In this chapter, we will discuss the ASA methodology in detail.

Abstract

Ensemble-based parameter estimation for a climate model is emerging as an important topic in climate research. For a complex system as a coupled ocean-atmosphere general circulation model, the sensitivity and response of a model variable to a model parameter could vary spatially and temporally. Here, an adaptive spatial average (ASA) algorithm is proposed to increase the efficiency of parameter estimation. Refined from a previous spatial average method, the ASA uses the ensemble spread as the criterion for selecting “good” values from the spatially varying posterior estimated parameter values; the “good” values are then averaged to give the final global uniform posterior parameter. In comparison with existing methods, the ASA parameter estimation has a superior performance: faster convergence and enhanced signal-to-noise ratio.

4.1 Introduction

Parameter estimation using ensemble-based filter (Anderson, 2001) is emerging as a promising approach to optimize parameters in a complex model (Annan and Hargreaves, 2004; Hacker and Snyder, 2005; Annan et al., 2005 a & b; Ridgwall et al., 2007; Hacker and Snyder, 2005; Aksoy et al., 2006 a & b; Tong and Xue 2008 a & b; Nielsen-Gammon et. al., 2010; Hu et al., 2010; Zhang et al, 2012; Zhang, 2011 a & b; Wu et al., 2012 a & b; Liu et al. 2014). In parameter estimation in a complex system, such as a coupled ocean-atmosphere general circulation model (CGCM), one common issue is sampling error accumulation when a large number of observations are used to update a single-value parameter sequentially (Aksoy et al, 2006a). To address this issue, Aksoy et al (2006a) proposed a spatial updating technique that transforms a single-value parameter into a two-dimensional field and updates the field spatially, so that localization in filtering can limit the observational error accumulation. The final model parameter after each analysis has been derived in two methods. In the first method, the globally uniform parameter value is recovered using a spatial average of the entire spatially varying parameter field (denoted as SA, Aksoy et al, 2006a & b). In the second method, the spatially varying parameters are allowed to vary spatially after each analysis, in the so-called Geographically-dependent Parameter Optimization (denoted as GPO, see Wu et al., 2012, 2013).

Here, our objective is the recovery of the spatially uniform parameter value. We propose an average method called adaptive spatial average method (ASA). The ASA is refined from the SA method to increase the efficiency of parameter estimation. The ASA uses the ensemble spread as the criterion for selecting “good” parameter values from the spatially varying parameter estimation; these “good” values are then averaged to give the final posterior parameter. Liu et al. (2014) has recently shown some examples of successful ASA estimation in a CGCM. In this chapter, we will examine in detail the ASA methodology for parameter

estimation in a CGCM using ensemble-based filter. The e-folding solar penetration depth (SPD) is used as the major parameter for estimation in this study. We will show that, compared with the SA method and the GPO method, our proposed ASA produces a faster convergence rate for parameter estimation. The chapter is organized as follows. Section 4.2 briefly describes the parameter estimation scheme and the CGCM used in this study. Section 4.3 shows the model sensitivity to the parameter (SPD). Section 4.4 discusses the ASA method. The ASA method is compared with GOP method and SA method in section 4.5. A summary and further discussion are given in section 4.6.

4.2 Model and Method

4.2.1 Fast Ocean Atmosphere Model (FOAM)

Our model, the Fast Ocean Atmosphere Model (FOAM, Jacob, 1997) is a CGCM with an atmospheric component of a R15 (7.5° longitude, 4° latitude and 18 layers). The ocean component is a z-coordinate model with a resolution of 2.8° longitude, 1.4° latitude and 24 layers. Without flux adjustment, the fully coupled model has been run for over 6000 years with no apparent drift in tropical climate (Liu et al., 2007a). In spite of its low resolution, FOAM has a reasonable tropical climatology (Liu et al., 2003), ENSO variability (Liu et al., 2000), and Pacific decadal variability (Wu et al., 2003, Liu et al., 2007b).

4.2.2 Data Assimilation Scheme

We will use a particular EnKF scheme, the Ensemble adjustment filter (EAKF, Anderson, 2001, 2003) in this study. Model parameters will be estimated simultaneously with the state variables by augmenting state variables with model parameters (Banks, 1992a, b; Anderson, 2001).

The e-folding solar penetration depth (SPD) is used as the major testing parameter for estimation. Solar attenuation in the ocean is a function of the amount of biomass in the upper layers of the ocean (Smith and Baker, 1978; Ohlmann et al., 2000). Previous studies suggest that solar penetration can have a significant impact on the surface climate in a climate model (Schneider and Zhu, 1998; Nakamoto et al., 2001; Murtugudde et al., 2002; Ballabrera-Poy et al., 2007; Anderson et al. 2007). In particular, some modeling studies found that a deeper solar attenuation leads to warming in the tropical Pacific annual mean SST, which may then reduce the cold bias in the equatorial Pacific in a coupled ocean-atmosphere model (Murtugudde et al., 2002; Ballabrera-Poy et al., 2007; Anderson et al. 2007).

Following Murtugudde et al. (2002), the downward solar radiation $I(z)$, at depth of z in FOAM is calculated as

$$I(z) = I(0)\gamma e^{\left(-\frac{z}{h}\right)}$$

where $I(0)$ is the total incident solar radiation at sea surface and $\gamma=0.47$ (Frouin et al., 1989) represents the fraction of total solar radiation in the photosynthetically available radiation band (wavelengths from 380 to 700nm). The remaining fraction of solar radiance is fully absorbed in the top model layer of 20 meters. The h is the e-folding depth of the solar penetration depth (SPD), which will be estimated in our experiments. In the real world, the SPD can be treated as a state variable, too, because it can be calibrated using the remote sensing observation of ocean color. Here, however, it is treated as a model parameter that will be estimated using conventional observation of sea surface temperature (SST) and salinity (SSS).

In this chapter, we assume the “truth” SPD has a globally uniform value of 17-m, and the truth simulation is performed with this SPD. The first guess of SPD is assumed 20-m with an uncertainty of 3-m (standard deviation). The observation for the assimilations are the monthly

mean SST and SSS, which are generated by adding a Gaussian white noise to the corresponding “truth” states at each grid point. The observational error scales (standard deviation) are 1°K for SST and 1psu for SSS. An ensemble size of 30 is used in all of our experiments. A 30-year simulation from the control truth run is used for the initialization of the ensemble, with the restart file of January 1st of each year used as the initial condition for each ensemble member. For state variable, the upper 8 layers of ocean temperature and salinity (0~235m) are updated by the observations. The Gaspari and Cohn (1999) covariance localization is used with the influence radius of 3 grid points horizontally for both state variables and the parameter SPD. To extract signal-dominant state-parameter covariance, the enhanceive parameter correction is applied (DAEPC, Zhang et al., 2012). Before the parameter estimation is activated, the data assimilation is performed in a “spin-up” period of 2 years in which only the state variables are estimated.

4.3 Model sensitivity with respect to solar penetration depth

We first investigate the model sensitivity to the parameter, solar penetration depth (SPD). Two types of parameter sensitivities need to be considered when DAEPC is used to improve the model climate. The first type is the sensitivity of the response of the model climatology to the change of the parameter; this sensitivity shows if the final model climate can be improved by tuning this specific parameter. The ocean surface climates of FOAM are significantly different between a deeper SPD (20-m) simulation and a shallow SPD (17-m) one, characterized by a warming of up to over 0.5°K in the tropical ocean and a cooling of up to -0.5°K in the subtropical ocean (see Fig. 3.1 in chapter 3).

The second type of sensitivity tests the model’s sensitivity to parameter uncertainty (represented, say, by the ensemble spread of the parameter) in the observational space at the

observational time interval; this sensitivity examines the possibility of reducing parameter uncertainty using the observations available. Furthermore, the model response to parameter uncertainty consists of linear and non-linear parts. Since the Kalman Filter framework is derived as the optimal analysis for a linear system, some features involving non-linear dependence may be regarded as noise for parameter estimation. Successful parameter estimation requires a signal-dominant state-parameter covariance, which is derived most favorably in a model whose state variables exhibit a strong linear dependence on model parameters (Aksoy et al., 2006 a, b).

An ensemble simulation starting from the same initial condition but using different values of the parameter SPD (i.e. an perturbed ensemble of parameters) demonstrates the second type of sensitivity (Fig.4.1). (Here, the parameter ensemble is constructed as a Gaussian distribution with the mean of 20-m and the standard deviation of 3-m). Since we will use the observations of monthly SST for parameter estimation, we will examine the ensemble response of the first month SST. The ensemble spread of the first month SST (monthly mean) represents the response of the model SST to the uncertainty of SPD in the observational space; the correlation coefficient between the SPD ensemble and the first month SST quantifies the linear part of the response. Fig.4.1 shows an overwhelmingly negative correlation between SST and SPD, implying predominantly a colder SST with a deeper SPD. This cooling is likely to be caused by the direct effect of solar penetration. Physically, a deeper SPD allows more solar radiation to penetrate below the surface layer, leaving less shortwave radiation heating the surface layer, and therefore cause surface cooling. The direct effect of solar penetration is dominant in the initial months in response to a sudden change of the SPD (Hokanson, 2006). One striking feature of the sensitivity is the strong variation with season and location. The SST ensemble spread is large and exhibits negative correlations in the summer hemisphere where the mixed layer is shallow and therefore

the SST is more sensitive to heat flux perturbations. Fig.4.1 is important for our parameter estimation, because it indicates the key regions for parameter estimation. The regions with large sensitivity and high correlation represent the regions of large linear model response to SPD. These regions have high signal/noise ratio and therefore are the regions where the observation of SST are most effective for parameter estimation. The rest of regions, which account for more than half of the grid points at each analysis step, are unlikely to provide significant information for parameter estimation.

4.4 The Adaptive Spatial Average scheme (ASA)

The sensitivity experiments in section 4.3 show that the model response to the parameter SPD varies significantly in both space and time. We speculate that neither GPO nor SA is most efficient for estimating the parameter. This follows that only the regions with large model-to-parameter linear response can provide state-parameter covariance with high signal/noise ratio for parameter estimation. Figure implies that the state-parameter covariance is insignificant over about half of the grid points at a time and in about half of the year at a given grid point. Therefore, for the purpose of parameter estimation, the estimations are not useful for more than half of the time at a given grid, and the estimations are not useful for more than half of the grids in the basin for a given observation time. Therefore, both SA and GPO are not the most efficient methods to estimate the parameter SPD, as will be shown below.

Here we refine the SA method to the Adaptive Spatial Average (ASA) method, to increase the efficiency of parameter estimation. In SA, the final spatially uniform parameter is estimated as the average of all the spatially different posteriors, each derived at a grid point using localization. The ASA is based on the idea that a parameter estimation, which will be derived

from an average of spatially different posteriors, should be more accurate if it only includes average those posteriors of smaller uncertainties (i.e. errors). For practical applications where the truth parameter, and therefore, the parameter error, is unknown, we can consider the ensemble spread as a representation of the error, as in traditional application of ensemble filtering to state variables (e.g. Evensen, 2007). (We will return to this point later). Therefore, the ensemble spread can be considered as the indicator of the quality of each posterior parameter values and a higher quality posterior has a smaller ensemble spread. The ASA will only retain those high quality values for the final averaging to derive the value for the spatially uniform parameter. This average value of high quality values should have smaller error than the average value of averaging all the values as in SA, which include the high quality as well as low quality values. A preliminary theoretical analysis of this point is given in the appendix.

A posterior value is “good” if its ensemble spread is relatively small among all the posteriors estimated at all the grid points. In practice, we use a threshold of the spread ratio between the posterior and the prior to judge the quality of the posterior and a posterior with a spread ratio below the threshold is considered a “good” posterior to be included for the final spatial average. (It should be noted that the ensemble spread of the prior is spatially uniform over the globe. Therefore, this spread ratio of the posterior over prior does not affect the relative magnitude of the posterior.) The speed of the decrease of the parameter uncertainty depends greatly on the magnitude of the signal. Initially, the ASA can use a small ratio as the threshold because the initial parameter uncertainty is large and the response magnitude (signal) is large. The threshold will be increased during the simulation with the decrease of the parameter uncertainty. The ASA is applied every few EnKF analysis cycles to obtain sufficient numbers of “good” parameter posterior values. The ASA therefore differs from the SA of Aksoy et al

(2006a), in which the spatial average is performed every EnKF analysis cycle and on all grid points. A conditional covariance inflation technique (CCI) as in Aksoy et al. (2006b) is also employed here on parameter ensemble after each ASA step to avoid the filter divergence for parameter estimation. The CCI inflates the parameter ensemble back to a predefined minimum value when necessary. The predefined minimum value is also the final uncertainty target for the estimated parameter.

4.5 Comparison of ASA with GPO and SA

We now compare ASA with SA and GPO schemes in FOAM. Two sets of experiments of parameter estimation are performed using observations of monthly SST and SST at every grid point. The first set of experiments (EXP-1a and EXP-1b) use the GPO scheme and confirm that the parameter ensemble spread is a good index for the parameter uncertainty (Figs. 4.2, 4.3). The second sets of experiments (EXP-2a and EXP-2b, Figs. 4.4, 4.5) compare the parameter estimations between SA and ASA schemes. The details of experimental setting are shown in table 4.1.

4.5.1 The assimilations with GPO scheme

Both EXP-1a and EXP-1b use the GPO scheme but with different observations. EXP-1b uses regular observations that consist of the “truth” plus noise. EXP-1a, called perfect observation experiment, uses the “truth” from control as the observations but nevertheless treats it as having the same uncertainty scale as in EXP-1b. For these two GPO experiments, neither EXP-1a nor EXP-1b is able to produce good parameter estimation, if only the monthly SST and SSS data are assimilated. Therefore, we are forced to also assimilate daily atmosphere wind (U, V) and temperature (T) with the error scale of 1 m/s and 1K, respectively; the observational error

scales for SST and SSS are also forced to be reduced from 1°K and 1 psu to 0.5°K and 0.5 psu, respectively. The initial SPD error is also reduced from 3-m to 1-m.

As speculated, the spatial pattern of the RMSE of SPD in EXP-1a is very consistent with the ensemble spread after 20 years of simulation (Figs. 4.2a, b). There are some regions of low uncertainty of SPD in different ocean basins. A further study shows that the low uncertainty in the mid-latitude North Pacific and North Atlantic are related to the large model sensitivity to SPD during the boreal summer (Fig.4.1b) and fall (Fig.4.1c); the low uncertainty in the eastern South Pacific, western equatorial Pacific, South Atlantic and southern Indian Ocean are partly related to the large sensitivity of the model SST to SPD in the austral fall (Fig.4.1a) and summer (Fig.4.1d). The high positive correlation between the parameter uncertainty and its ensemble spread can be seen more clearly in the scatter plot, for example, at the simulation year of 40 (Fig.4.3a). The RMSE of SPD estimation and its ensemble spread show a strong positive linear correlation with only modest spread residual. The estimate values are closer to the truth when the ensemble spread is small, except for the case of very small ensemble spread (~ 0.3 in Fig.4.3a). The positive correlation between the posterior error and ensemble spread supports our speculation before that the ensemble spread can be used to represent the estimation error or uncertainty. Furthermore, it is clear that a spatial average will decrease the parameter error because the average reduces the part of parameter uncertainty that is spatially independent (see eqn. (A4.4 in the appendix). The error of SPD can be further reduced by using only the posterior values with smaller ensemble spread for average (Fig4.3b), as hypothesized for the ASA. The error of SPD is reduced to 0.40-m when the posterior values of SPD over all the global grid points are averaged in EXP-1a (after 40 years of assimilation), compared with the global mean RMSE of SPD of 0.6-m (first RMSE and then global average); this error is decreased to 0.2-m

and 0.1-m when the top 50% and 20% of grid points of smallest ensemble spread are averaged, respectively. When the ensemble spread is at its smallest values, the estimated values suffer from an overshoot, i.e. the parameter error becomes negative. This phenomenon also occurs in chapter 3 when the similar observation coverage is applied, i.e. U, V and T for the atmosphere and SST and SSS for the ocean. The reason for the overshoot will be discussed in a future study.

The positive correlation between the parameter uncertainty and parameter RMSE, however, is disrupted significantly when the regular observation (“truth” plus noise) is used as in EXP-1b. Now, the spatial pattern of the parameter ensemble spread (Fig.4.2d) remains similar to that in EXP-1a (Fig.4.2b), but the pattern of the SPD uncertainty (Fig.4.2c) become very noisy. This occurs because the parameter updating using EnKF also introduces observational errors into the SPD posterior, which is equivalent to adding random noise onto the parameter posterior of EXP-1a. This noise leads to a decrease of the consistence between the SPD uncertainty and its ensemble spread. The distortion on the correlation is seen clearly in the scatter plot Fig.4.3c, where the error value of SPD and its ensemble spread of EXP-1b shows a very weak linear relationship with a much-enhanced residual variance. Nevertheless, this correlation is still significant at the 99% level. Furthermore, since the uncertainty associated with the observation errors is spatially independent, it can be reduced dramatically using a spatial average. Indeed, the averaging values of SPD are very similar for EXP-1a and EXP-1b (Figs.4.3b vs. d), although the estimated values of SPD are much more noisy in EXP-1b than in EXP-1a.

Overall, the consistency between the parameter uncertainty and its ensemble spread indicates that the parameter ensemble spread can be used as a good index for the uncertainty of the parameter value and therefore can be used as the criteria for selecting “good” posteriors for averaging. A spatial average of those “good” posteriors tends to give a better final estimation.

4.5.2 Comparison between SA and ASA

As discussed regarding EXP-1a, 1b, and in the appendix, the uncertainty of the parameter posterior can be reduced using spatial average. The ASA and SA are applied in EXP-2a and EXP-2b, respectively. A predefined minimum ensemble spread value of 0.3-m for the CCI is applied in the EXP-2 (s). Unlike the GPO experiments above, now, the error of SPD is reduced dramatically in both EXP-2a and EXP-2b even only with monthly mean SST and SSS observations (Fig. 4.4a), implying an increased robustness of parameter estimation using spatial average.

Based on the ensemble sensitivity shown in Fig.4.1, we apply the ASA every 6 analysis cycles (6 months) in EXP-2a with an initial threshold of 0.68. To prevent the degeneration case of too few “good” values, the threshold increases by 0.1 until it reaches 0.98 whenever the total number of “good” values is smaller than a given number, here set as 400. The ASA picks different grids at different times for averaging. The number of grid points of “good” values also varies temporally in the range of 400~4000, which is around 2~40% of total ocean grids (Fig.4.4b). The ensemble spread of SPD initially decreases much faster than its real uncertainty (Fig.4.4a), reaching the minimum parameter ensemble spread of 0.3-m in 5 simulation years. Although this ensemble spread (0.3) is smaller than the real error in years of 5-20, the SPD continues to converge to its “truth”. The SPD error in EXP-2a is decreased from 3-m to 0.3-m (the estimating goal) in 20 years (Fig. 4.4a).

During the assimilation cycle, the ensemble spread still remains positively correlated with the estimation errors among different points, albeit with a substantial spread (as discussed for Exp.1b in Fig.4.3b). This can be seen in the two examples of scatter plots of SPD after the first and fifth spatial updating cycles in Fig.4.5a and 4.5b, respectively. The ASA produces a

good SPD estimation by averaging only a moderate number of “good” values (200-2000) once the threshold (the uncertainty ratios between the posterior and prior) is selected appropriately. This can be seen in Figs.4.5c and Fig.4.5d, which shows the number of “good” values and the average of these “good” values respectively, as functions of the threshold in ASA for the first 5 assimilation cycles. For example, for the first assimilation cycle, the average of SPD is 18.5 with the threshold of 0.8 and the number of “good” value of ~400; the average SPD is 17.6 with the threshold of 0.65 and the number of “good” value of ~1000. If the threshold is too small, too few values are defined as “good” values. This will lead to a too small sample size and large sampling error, such that ASA no longer produces good results (Fig.4.5b & d).

The final estimation also depends on the minimum ensemble spread specified in CCI. The error of the estimated SPD seems to saturate at the equilibrium level of ~0.2-m error in ~30 years in EXP-2a if the minimum parameter ensemble spread remains at 0.3-m. This minimum ensemble spread can be decreased afterwards to yield more accurate estimation. The ASA estimation is repeated from year 31 to year 47 but now with the minimum parameter ensemble spread reduced from 0.3-m to 0.2-m; now the SPD error further decreases from 0.2-m to ~0.1-m (Fig.4.4a, green lines). In this case, a reduced minimum ensemble spread further improves the final convergence of the parameter estimation.

In comparison with the ASA (in EXP-2a), the spatial average using all the grid points in SA (EXP-2b) shows a considerably slower convergence in the SPD estimation, with the SPD error barely reaching 0.3-m after 47-years of assimilation (red lines, Fig.4.4a). Similar to the ASA, the ensemble spread of SPD in SA also decreases much faster than its real error scale. The CCI with the minimum parameter ensemble spread of 0.3 prevents the filter divergence of the parameter estimation. In the mean time, the evolution of estimation SPD in SA is more stable

than in ASA because more grids and in turn a bigger sample size in the former than the latter. Overall, ASA demonstrates a faster convergence rate than SA for SPD estimation because the former uses only “good” values for averaging.

4.6. Summary and Discussions

Refining the Spatial Average scheme (SA), we proposed the Adaptive Spatial Average scheme (ASA) to improve the efficiency of the parameter estimation in a complex system, such as a CGCM. The ASA is explored in the twin experiment framework in FOAM, where the biased parameter (SPD) is the only model error source. The e-folding scale of the solar penetrating depth is used as the biased parameter for estimation. Sensitivity experiments show that the response of the FOAM to the parameter uncertainty varies spatially and temporally. The ASA is demonstrated to increase the efficiency of parameter estimation significantly over previous assimilation techniques such as the SA (Aksoy et al., 2006a) and geographic dependent parameter optimization (GPO) (Wu et al, 2012a).

The ASA uses the posterior ensemble spread as the criterion to select the “good” values from the spatial updating posterior parameter values and only use the “good” values for the averaging to yield the globally uniform posterior. In comparison with the SA scheme, the ASA produces a faster convergence for parameter estimation. The faster convergence of ASA than SA is robust in other settings, as seen in two additional pairs of experiments the same as EXP-2a & b, except for the observational interval of 10 days (EXP-3a & b) and 1 day (EXP-4a & b), respectively (table 4.1). When the observational interval is shortened, the model response to the parameter uncertainty becomes more linear. However, the response amplitude still varies spatially and temporally (not shown). Therefore, ASA is still more suitable than SA. Similar to

EXP-2, both EXP-3 and EXP-4 show faster decreases of the SPD ensemble spread than its real uncertainty in the initial stage. The convergence time is also shortened for a shorter observational interval. In ASA, the SPD errors reach the objective uncertainty (0.3-m) in ~ 10 years (EXP-3a, Fig.4.6a) and ~ 5 years (EXP-4a, Fig.4.6b) of simulations, for the observational interval of 10 and 1 days, respectively, while, in SA, they take ~ 30 years (EXP-3b, Fig.4.6a) and ~ 10 years (EXP-4b, Fig.4.6b). It is noted that the estimated SPD in EXP-4 (Fig.4.6b) is less stable than that in EXP-2 or EXP-3 (Fig.4.3a, Fig.4.6a). The observational interval in EXP-4 is only 1 day, while the decorrelation time scale of SST is a few months. This results in the accumulation of sampling error because the model SST does not have the time to respond before another observation is added. The accumulation of sampling error causes poor parameter estimation compared to the other experiments. Furthermore, the instability of the estimated parameter in Fig.4.6b could become worse as the total assimilation time increases. We could increase the assimilation time interval for parameter estimation to reduce the instability of parameter estimation.

The ASA is designed to deal with the spatially and temporally varying feature of model response to parameter in CGCM. As pointed out by one reviewer, for SPD, SST shows little sensitivity to the parameter perturbation in about half of the world ocean (Figs.4.1a-d). One may speculate that our experiments for the estimation of SPD are too peculiar. The SA is inferior to ASA because the posteriors in these regions of little sensitivity are subject to too large a noise (with little response signal) and therefore contaminate the SA estimation seriously. To clarify this, it will be desirable to test the estimation for a parameter that has a more spatially uniform response sensitivity. Therefore, we repeated the estimation for two other parameters m_d and m_q in chapter 3. The m_d and m_q are artificial multipliers to the momentum and latent heat fluxes between the ocean and atmosphere, respectively, with 1 as the default truth model value. The

model SST sensitivity to either parameter is more uniform than for SPD (not shown). Our experiments EXP-5a and b and EXP-6a and b use the same experimental setting as EXP-2a and b except for estimating the imperfect parameter m_d and m_q , respectively (table 4.1, Fig.4.7). Both EXP-5a and EXP-6a show faster decreases of the parameter errors than EXP-5b and EXP-6b. The m_d reaches the objective uncertainty of 0.04 (set by the minimum ensemble spread specified in CCI) in ~ 10 years with ASA but in more than 30 years of assimilation with SA (Fig.4.7a). Similarly, the m_q reaches the objective uncertainty of 0.04 in ~ 25 years with ASA but in more than 40 years of assimilation with SA (Fig.4.7b). Therefore, the improvement of ASA over SA is valid for more general cases than the SPD.

The ASA has also been shown successful for the estimation of multiple parameters (Liu et al., 2014). Therefore, we believe that the ASA method is well suited for the estimation of those parameters with a globally uniform feature in CGCM. The estimation of a spatially varying parameter in CGCM, however, remains to be further studied.

4.7 Appendix: Preliminary Theoretical Consideration for ASA

Here, we will discuss the SA and ASA from a more quantitative perspective. When we implement the spatial updating in ensemble-based parameter estimation, we obtain a spatially varying parameter posterior field. The posterior errors at different locations are correlated because the parameter priors are identical for the entire field. To quantify the effect of spatial averaging, we can separate the posterior errors into two independent components: one linearly dependent on the parameter prior error and the other uncorrelated with the first one.

In EnKF, the covariance(s) between the parameter and the model forecasts in observational space are used directly to update parameter in exactly the same manner as for the

state variables. When we use a forecast x^f and an observation x^o to update a parameter β , the $(\sigma_\beta^2)_a$ of a parameter posterior can be written as

$$(\sigma_\beta^2)_a = \sigma_\beta^2(1 - \theta) \quad (\text{A4.1})$$

where $\theta = \frac{\rho^2 \sigma_x^2}{(\sigma_x^2 + R)}$ with $0 \leq \theta < 1$. Here the σ_x^2, R are the error scales (variances) of x^f and x^o , respectively; ρ is the correlation coefficient between forecast x^f and parameter prior. The uncertainty of parameter posterior decreases with the increase of θ . The ratio between parameter posterior uncertainty and prior uncertainty

$$\frac{(\sigma_\beta^2)_a}{\sigma_\beta^2} \equiv r = 1 - \theta.$$

In EnKF, $(\sigma_\beta^2)_a$ and σ_β^2 are represented by the variance of parameter posterior and prior ensemble, respectively. So the r is the ratio between the posterior and the prior ensemble spread. For a spatial updating, different location has different r . The ASA uses the r as index to select the “good” values from a posterior field.

The parameter posterior error of ε_β^a originates from different sources: x^f, x^o and β^f , and can be written into two parts based on the correlation relationships among the error sources

$$\varepsilon_\beta^a = \sigma_\beta N_\beta^f (1 - \theta) + \sigma_\beta \sqrt{\theta - \theta^2} N_x^b \quad (\text{A4.2})$$

where N_β^f and N_x^b are independent white noise with the scale of 1. The two terms on the right hand side of (A4.2) represent two independent components of the total uncertainty (error) of a posterior value for any given σ_β and θ . The 1st term linearly depends on the error of parameter prior of $(\sigma_\beta N_\beta^f)$, while the 2nd term is uncorrelated with the error of parameter prior. The 2nd term is produced by the errors from observations, initial conditions and the nonlinear part of model response to the parameter prior. The 1st term is dominant when θ is close to 0 and the uncertainty

of posterior is close to the uncertainty of parameter prior. The second term become primary when θ is close to 1 and the uncertainty of posterior is much smaller than the uncertainty of parameter prior (Fig. 4.8).

For a spatial updating, we can rewrite (A4.2) into a spatially varying field

$$\varepsilon_{\beta,i}^a = \sigma_{\beta} N_{\beta}^f (1 - \theta_i) + \sigma_{\beta} \sqrt{\theta_i - \theta_i^2} N_{x,i}^b \quad (\text{A4.3})$$

where $i = 1, 2, \dots, N$ indicate the locations. The first term on the right hand side is all linearly dependent among different locations, while the second term on the right hand side can be regard as independent among different locations when the posterior values are widely distributed over a large domain. For a spatial average, the two terms have opposite changes. Averaging the β_i^a to obtain a single-value parameter, the posterior error is

$$\overline{\varepsilon_{\beta,i}^a} = \frac{\sigma_{\beta}}{M} \sum_i N_{\beta}^f (1 - \theta_i) + \frac{\sigma_{\beta}}{M} \sum_i \left(\sqrt{\theta_i - \theta_i^2} N_{x,i}^b \right) \quad (\text{A4.4})$$

We now discuss the two terms on the right hand side of (A4.4) one by one, regarding the difference between SA and ASA. The first term is linearly dependent on the parameter prior error (N_{β}^f), therefore its scale mainly affect by the distribution of θ_i but not the averaging sample size of M . The first term can be discussed conveniently by assuming a uniform distribution $[\theta_{\min} \quad \theta_{\max}]$ for θ_i . The SA scheme (Aksoy et al 2006a) averages all posterior values over the entire domain. This term becomes $\sigma_{\beta} (1 - \frac{\theta_{\min} + \theta_{\max}}{2}) N_{\beta}^f$. The ASA sets a threshold θ_{th} ($\theta_{\min} \leq \theta_{th} \leq \theta_{\max}$) to remove the values with $\theta_i < \theta_{th}$ from the average pool such that this term becomes $\sigma_{\beta} (1 - \frac{\theta_{th} + \theta_{\max}}{2}) N_{\beta}^f$, which is smaller than that using the SA scheme when the difference between θ_{\min} and θ_{\max} is large and θ_{th} is significantly greater than θ_{\min} . When

$\theta_{th} = \theta_{min}$, the ASA recovers to the SA. When $\theta_{th} = \theta_{max}$, the ASA just picks the posterior value with the “best” posterior, i.e. the minimum analysis error.

The second term on the right hand side of (A4.4) decreases with the increase of the average sample size of M because the $N_{x,i}^b$ are independent among different sites. Therefore the second term in ASA is larger than that in SA because ASA uses a smaller M than SA. However, when the number of average values (M) is sufficiently large, the second term for both SA and ASA is smaller than the first term (unless the $\theta_i(s)$ are all close to 1), and therefore has limited impact on the total error. When the $\theta_i(s)$ are all close to 1, the first term is trivial comparing with the second term before average (see (A4.2) and Fig.4.8); but this rarely happens for parameter estimation with EnKF in a complex system like CGCM, because it would require $\rho^2 \approx 1, \sigma_x^2 \gg R$. The θ_{min} is usually close to 0, especially when the parameter is nearly converging.

The ASA can reduce the error related to the parameter prior error in spite of a reduced the averaging sample size, because “good” posteriors are used which have sufficiently large θ_i . The ASA produces a better analysis of β than SA after averaging the same posterior field when the θ_{th} is significantly smaller than the θ_{max} when they average the same posterior field. In summary, the SA reduces the errors related to the observations and forecasts. These errors are uncorrelated between different locations. The ASA scheme enhances the signal during the averaging by filtering out the region with weak signal or no signal. Therefore the ASA can produce a faster convergence than the SA (see Fig.4.4a, Fig.4.6 and Fig.4.7).

It should be mentioned that the sampling error generated by the limited ensemble size is ignored in our derivation. The sampling error is not represented by the parameter ensemble, and therefore will not affect the derivation discussed previously. Usually both SA and ASA can reduce sampling error to a small term compared with two error terms in (A4.4) when the number

of average values (M) is sufficient. However, when an estimated parameter is near its convergence limit, the σ_β is small and the $\theta_i(s)$ are close to 0. Although the 2nd term at the right hand side of (A4.4) is much smaller than the 1st term (See Fig.4.8), the sampling error could be comparable with the other two error sources in the equation even after the averaging. Therefore, the sample size of averaging values (M) becomes critical. Parameter estimation under this condition is very dangerous, especially for the ASA method because of its smaller average sample size. We can avoid this situation by applying a CCI as in Aksoy et al. (2006b). The minimum ensemble spread of CCI maintains the uncertainty of a parameter (σ_β) at a reasonable level. That makes the sampling error relatively small compared to the other two error terms.

Table

Table 4.1 the experiment setting. The oceanic observations are SST and SSS; and atmospheric observations are T, U and V. EXP-1a uses the perfect observations (truth). EXP-5 a & b estimate the parameter of m_d and EXP-6 a & b estimate the parameter of m_q .

EXP	Method	Obs. (Ocn.; Atm.)	Parameter (Truth)	Initial guess/Truth/ uncertainty
1a &b	GPO	1 month; 1 day	SPD	18m/17m/1m for SPD
2a	ASA	1 month; N/A	N/A	20m/17m/3m for SPD
2b	SA	1 month; N/A	N/A	20m/17m/3m for SPD
3a	ASA	10 days; N/A	N/A	20m/17m/3m for SPD
3b	SA	10 days; N/A	N/A	20m/17m/3m for SPD
4a	ASA	1 day; N/A	N/A	20m/17m/3m for SPD
4b	SA	1 day; N/A	N/A	20m/17m/3m for SPD
5a	ASA	1 month	N/A	1.2/1.0/0.2 for m_d
5b	SA	1 month	N/A	1.2/1.0/0.2 for m_d
6a	ASA	1 month	N/A	1.2/1.0/0.2 for m_q
6b	SA	1 month	N/A	1.2/1.0/0.2 for m_q

Figures

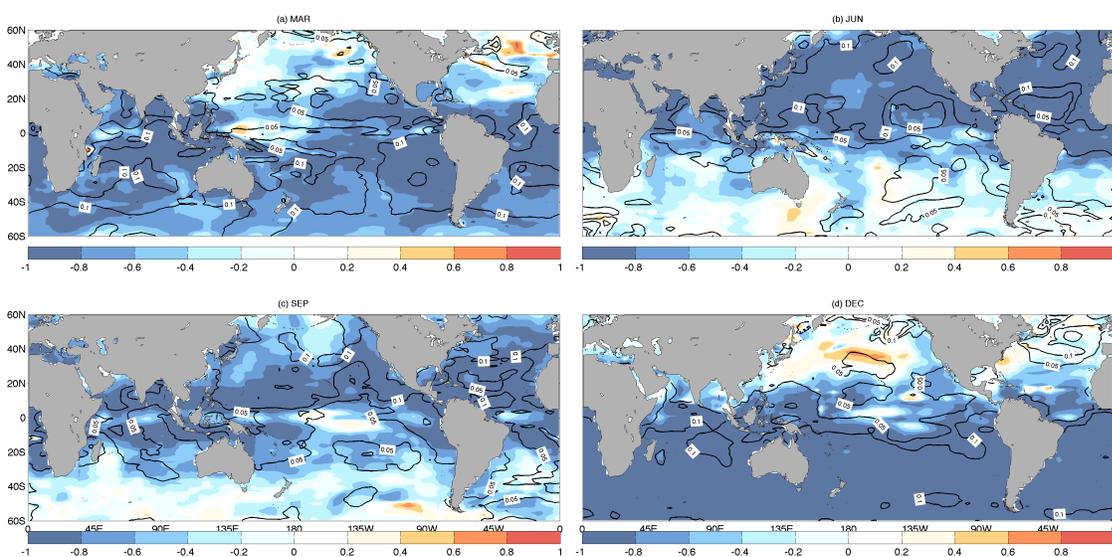


Fig. 4.1. The model monthly SST response to 3-m SPD uncertainty at different month for (a) March, (b) June, (c) September and (d) December. The shading represents the correlation coefficient between the SPD ensemble and the 1st month monthly SST response while the contours represent the magnitude of the monthly SST response (ensemble spread). A 30-member ensemble simulation that starts from the same initial condition but use different values of the parameter SPD. The SPD ensemble is constructed as a Gaussian distribution with the mean of 20-m and the standard deviation of 3-m. We integrate the model from the beginning of each month to the end of the month to obtain the monthly mean response.

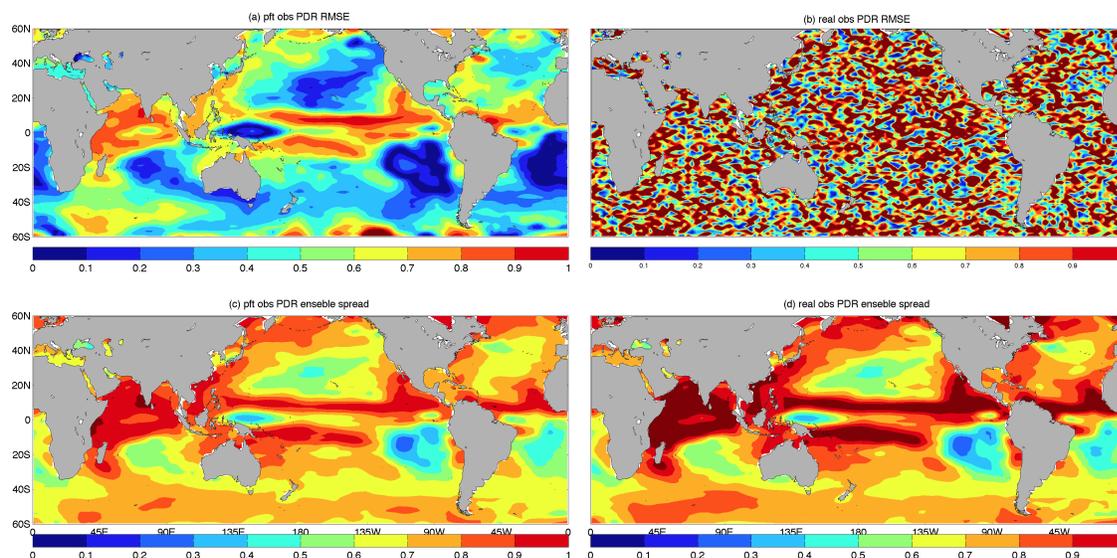


Fig. 4.2. Solar penetration depths estimated using DAEPC with the GPO method. The total ensemble size is 30. Panel (a) and (c) are the spatial distribution of parameter error values and parameter ensemble spreads after 20 years simulation for the perfect observation experiment. Panel (b) and (d) are the parameter error values and parameter ensemble spreads after 20 years simulation for regular observation experiments.

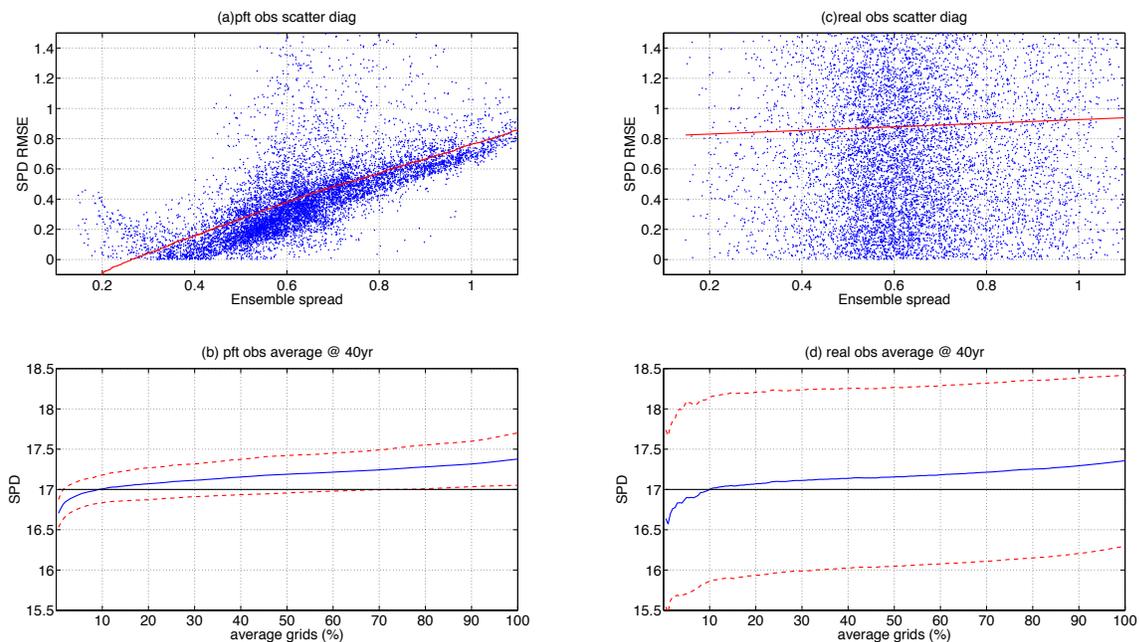


Fig. 4.3. The estimated SPD after 40-year simulations using DAEPC with the GPO method.

Panels (a) and (b) are for EXP-1a using perfect observations: (a) the scatter diagram between SPD error values and ensemble spreads. The red line is the regression line. (b) The blue line is the averaging value of SPD using top percent grids (with smallest ensemble spread) and the red dish lines represent 1-standard deviation of the averaging values. The black dish line is the “truth”.

Panels (c) and (d) are the same as (a) and (b) but for EXP-1b using regular observations.

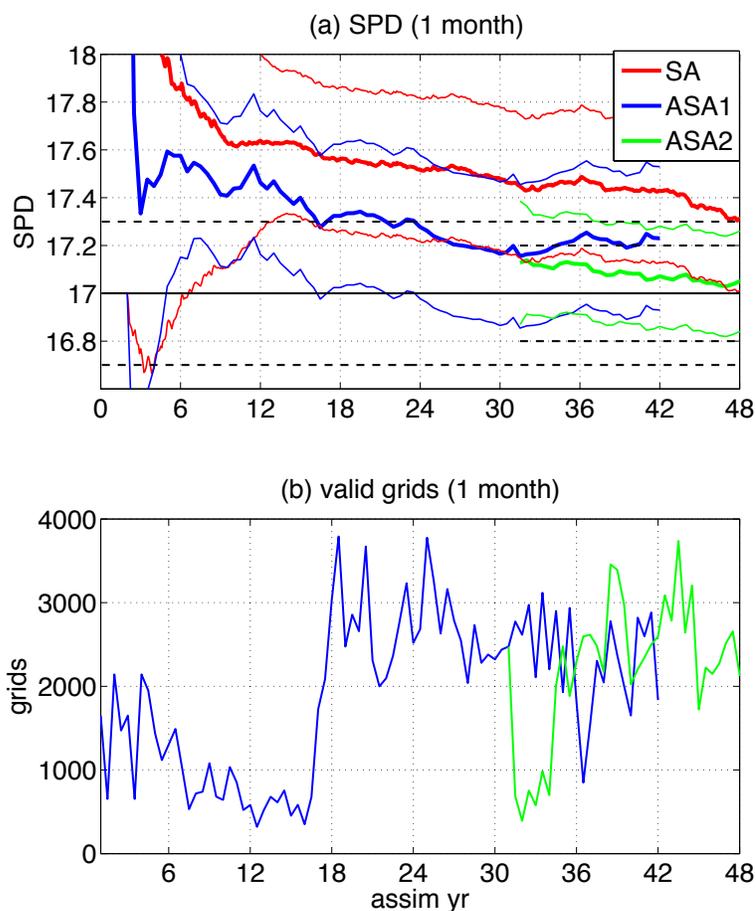


Fig. 4.4. The Estimated SPD using DAEPC with the ASA (EXP-2a) and SA (EXP-2b).

(a) Temporal evolution of parameter error (thick lines) and 1-standard deviation of ensemble spread (thin lines). The red lines are for EXP-2b and the blue lines are for EXP-2a the green lines are also for EXP-2a but with a reduced minimum parameter ensemble spread of 0.2 for the year 31~47. The black solid line is the “truth” and the black dish lines are the minimum parameter ensemble spreads (uncertainty goals) for the experiments. (b) temporal evolution of total numbers of grids used for average in ASA.

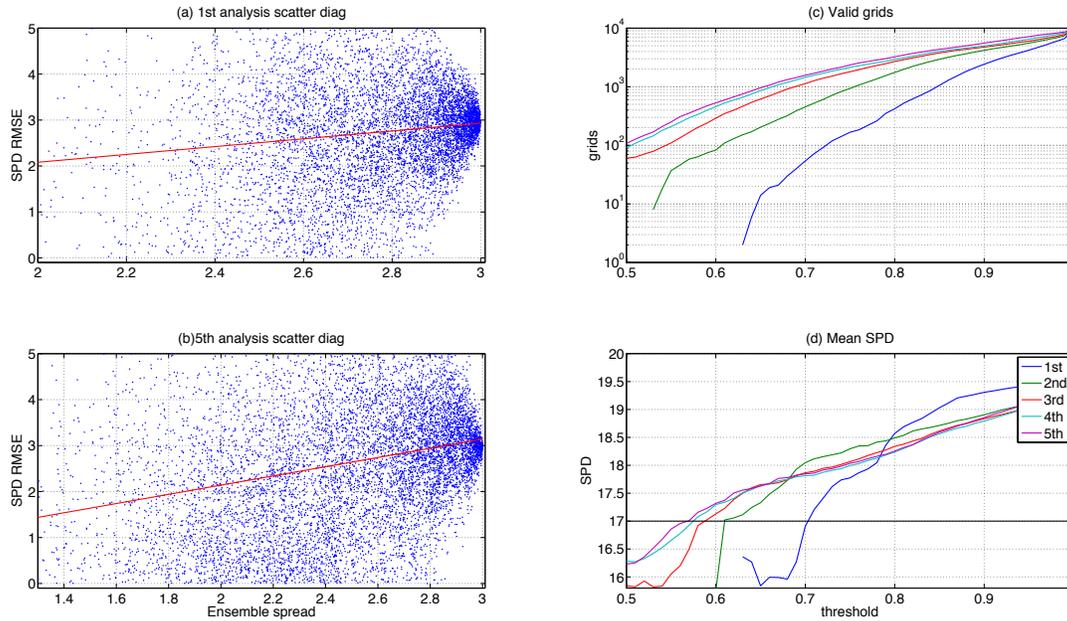


Fig.4. 5. (a) the scatter diagram between SPD error values and its ensemble spreads for EXP-2a after the 1st analysis cycle of parameter updating. The red line is the regression line. (b) is the same as (a) but for after the 5th analysis cycle.

(c) the numbers of “good” grids (values) for the 1-5 analysis cycles of EXP-2a using ASA with different threshold. The blue line is for the 1st analysis cycle, the green line is for the 2nd, the red line is for 3rd, the cyan line is for 4th and the magenta is for 5th.

(d) the mean SPD values of the “good” grids from (c) respectively.

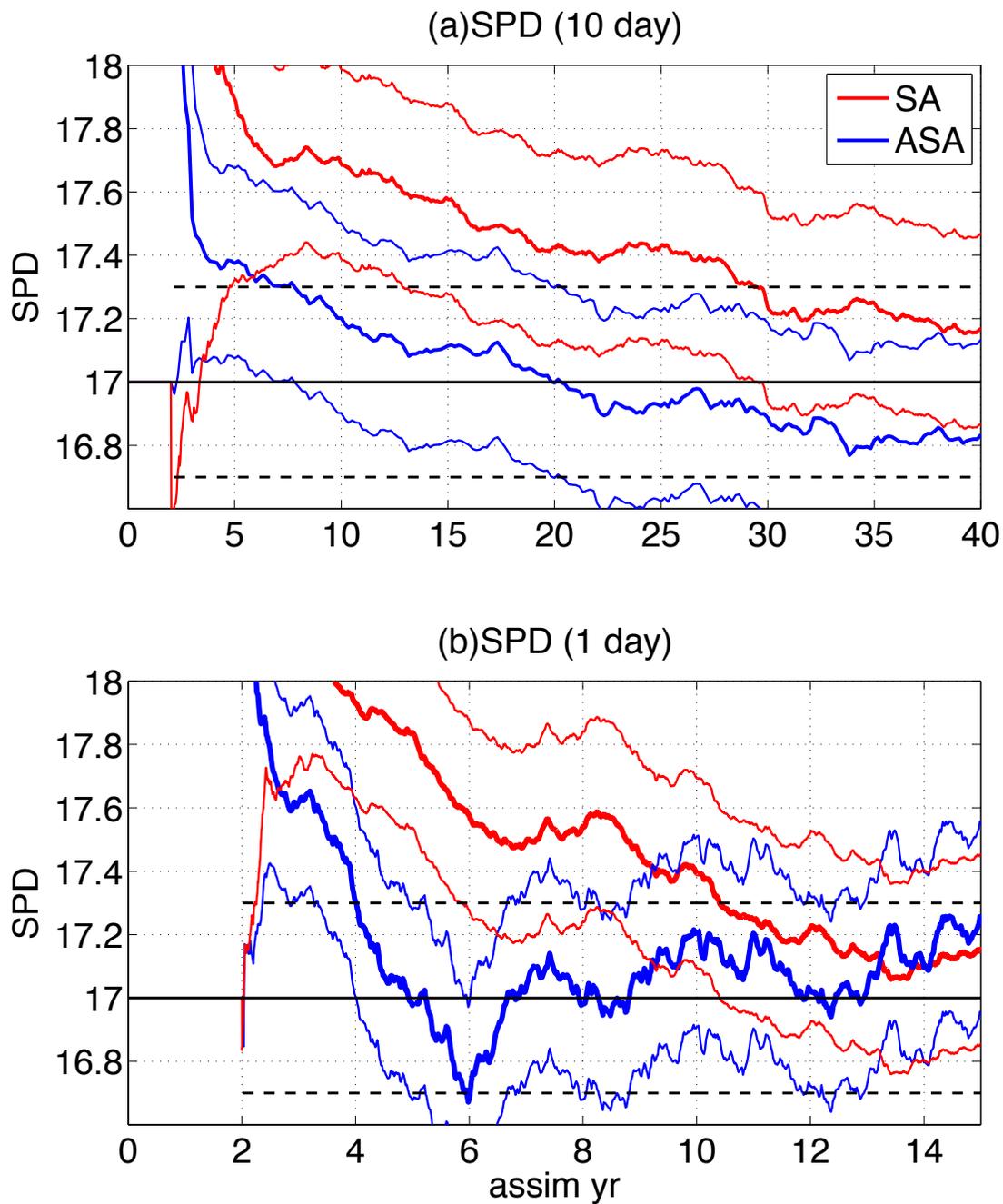


Fig. 4.6. (a) the temporal evolution of SPD (thick lines) and 1-standard deviation of ensemble spread (thin lines) for EXP-3. The red lines are for EXP-3b and the blue lines are for EXP-3a. The black solid line is the “truth” and the black dish lines are the minimum parameter ensemble spreads (uncertainty goals) for the experiments.

(b) is same as (a) but for EXP-4.

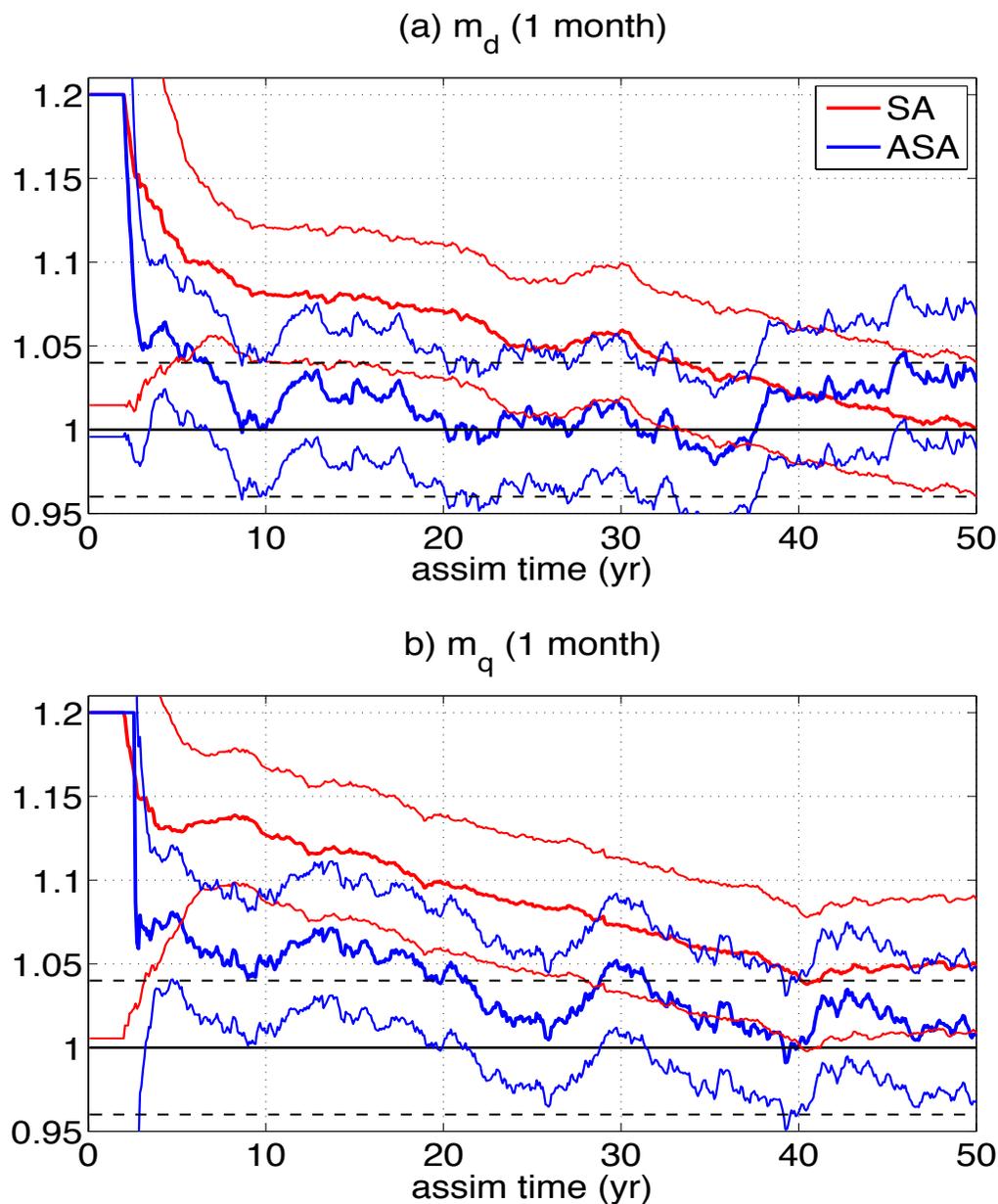


Fig. 4.7. (a) the temporal evolution of m_d (thick lines) and 1-standard deviation of ensemble spread (thin lines) for EXP-5. The red lines are for EXP-5b and the blue lines are for EXP-5a. The black solid line is the “truth” and the black dash lines are the minimum parameter ensemble spreads (uncertainty goals) for the experiments.

(b) is same as (a) but for the temporal evolution of m_q for EXP-6.

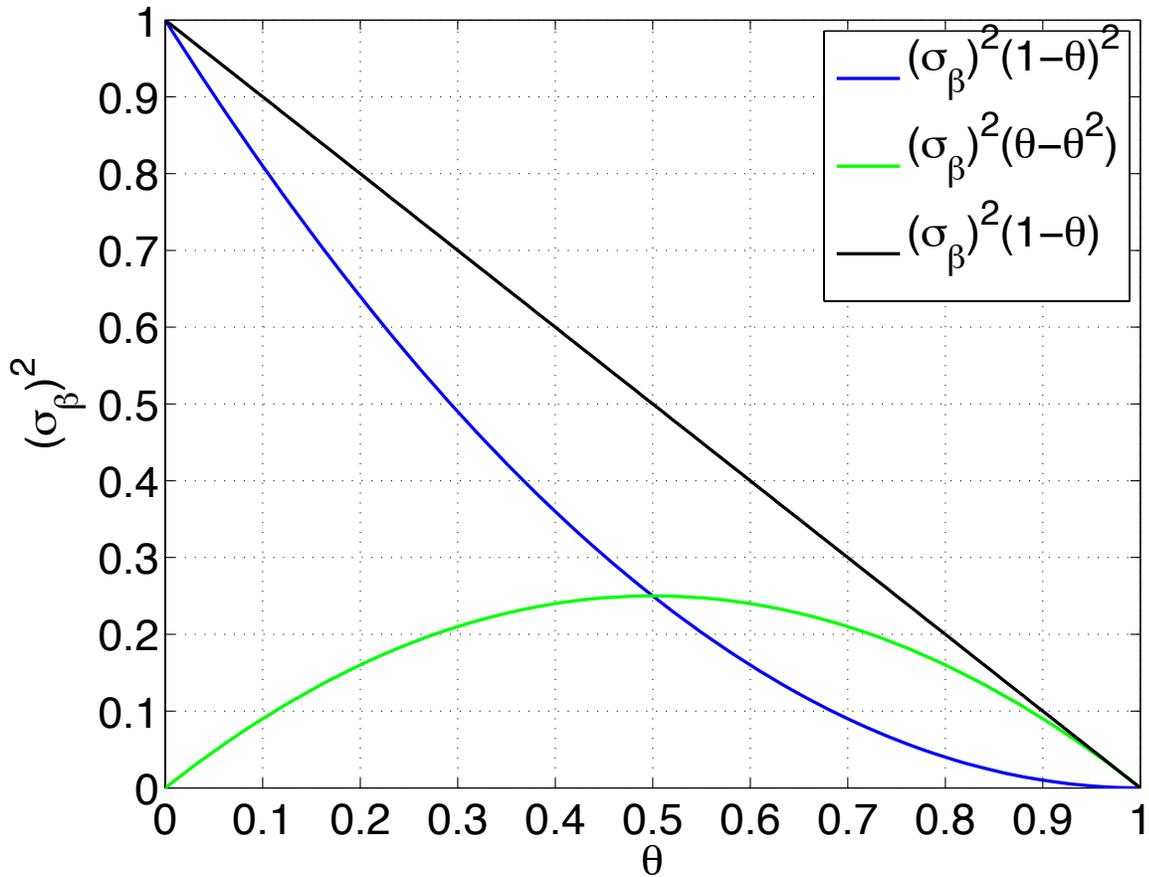


Fig. 4.8. The scale (variance) of each term in equation (A2). The blue curve represents the scale of the 1st ($\sigma_\beta N_\beta^f(1-\theta)$) at the right hand side of the equation, which is related to the error of the parameter prior; the green curve represents the scale of the 2nd term ($\sigma_\beta \sqrt{\theta - \theta^2} N_x^b$) at the right hand side of equation, which is related to the uncertainties of the observation and forecast, but unrelated to the parameter uncertainty. The black curve is the scale of the total error (ε_β^a).

Chapter 5 Summary and Future work

5.1 Summary

Coupled ocean-atmosphere general circulation models (CGCM) are very useful tools for climate study and climate prediction. The uncertainty of model parameters is one of the important sources of model bias that cause the model climate to drift from real world. The traditional method of tuning parameters through trial-and-error sensitivity experiments is subjective and very inefficient. In this study we apply a novel strategy for systematic parameter estimation in a CGCM using ensemble based parameter estimation (Anderson, 2001).

We developed a random subgrouping scale for deterministic ensemble based filter to improve the ability handling nonlinearity. The ensemble based parameter estimation is based on the data assimilation skill of the ensemble-based filter (Evensen, 1994, 2007; Burger et al., 1998; Houtekamer and Mitchell, 1998; Anderson, 2001, 2003; Bishop et al. 2001, Whitaker and Hamill 2002; Tippett et al 2003; Houtekamer et al., 2005; Sakov and Oke 2007). The ensemble-based filter is a linear filter, although it could be used in the nonlinear regime. The ability of an ensemble-based filter handling nonlinearities is very important for ensemble based parameter estimation because the model response to a parameter uncertainty could be very nonlinear. Based on the methods of updating analysis ensemble, ensemble-based filters can be divided into two categories: stochastic and deterministic. A deterministic filter avoids the sampling error introduced by perturbed observations and tends to generate better analyses than a stochastic filter when applied to a linear model, especially for a small ensemble size (~10-20) (Whitaker and Hamill 2002, Evensen 2003, Anderson 2010). However, a deterministic filter suffers from the problem of generating extreme outliers in the ensembles produced in a nonlinear system,

especially when the ensemble size is large (Lawson and Hansen, 2004; Anderson, 2010). To address the problem of extreme outliers, we developed a new random subgrouping technique for the deterministic filter. The random subgrouping technique randomly divides the full ensemble into sub-ensembles of equal size for each observation at each analysis step. All sub-ensembles are updated independently using deterministic filter algebra. Compared to regular stochastic and deterministic filters, the new technique significantly improves performance of data assimilation in the Lorenz-63 systems (Lorenz, 1963), the Lorenz-96 system (Lorenz, 1996) and a global QG atmospheric model coupled to a slab ocean.

In a twin experiment framework, we successfully implemented a simultaneous estimation of model state variables and parameter with the assimilation of time evolution observations in a CGCM – the fast ocean atmosphere model (FOAM) (Jacob, 1997) for both single parameter estimation and multiple parameter estimation. In the case of the single parameter estimation, the error of the parameter (solar penetration depth, SPD) is reduced by over 90% after ~40 years of assimilation of the conventional observations of monthly sea surface temperature and salinity. The results of multiple-parameter estimation are less reliable than the single-parameter estimation when the same experimental setting is used. Assimilating additional observations of atmospheric data compromises the less reliability of multiple-parameter estimation. The errors of the parameters are reduced by 90% in ~8 years of assimilation. Finally, the improved parameters improve the analysis quality of state variables as well as the model climatology. With the optimized parameters, the bias of the climatology of SST is reduced by ~90%. Our study suggests the feasibility of the ensemble based parameter estimation in a CGCM.

We proposed an adaptive spatial average scheme (ASA) is to increase the efficiency of parameter estimation. The ASA is well suited for the estimation of those parameters with a

globally uniform feature in CGCM. When we estimate globally uniform parameters in FOAM, a spatial updating technique is applied to prevent the sampling error accumulation issue when a large number of observations are used to update a single-value parameter sequentially. Proposed by Aksoy et al (2006a), the spatial updating technique transforms a single-value parameter into a two-dimensional field and updates the field spatially, so that localization in filtering can limit the error accumulation. To recover the globally uniform feature of a parameter, the parameter values are recovered by using ASA on the spatially varying parameter field. For a complex system such as a CGCM, the sensitivity and response of a model variable to a model parameter may vary spatially and temporally. The area with stronger linear response to a parameter uncertainty can produce better estimation of the parameter. Refined from a previous spatial average method (Aksoy et al, 2006a & b), the ASA uses the ensemble spread as the criterion for selecting “good” values from the spatially varying posterior estimated parameter values; the “good” values are then averaged to give the final global uniform posterior parameter. In comparison with the spatial average, the ASA parameter estimation has a superior performance: faster convergence and enhanced signal-to-noise ratio.

5.2 Concluding remarks

This work is a pilot study of parameter optimization in CGCMs using ensemble based filter strategy. We focus on the study of feasibility and effectiveness of parameter optimization. This work has made new contributions on ensemble-based filter and ensemble based parameter estimation.

Firstly, we developed a random subgrouping ensemble based filter scheme to improve the filter performance in a nonlinear regime.

Secondly, we proposed a new average scheme- adaptive spatial average scheme to increase the efficiency of parameter estimation in complex system such as a CGCM.

Finally, we presented the first study of successful ensemble-based parameter estimation in a CGCM in a twin model framework, demonstrating the feasibility of parameter estimation in a CGCM.

5.3 Future work

Much further work remains. The ultimate goal is to develop an effective coupled data assimilation strategy for parameter optimization with ensemble based filter to improve CGCMs using the real observational data. All of our experiments of parameter estimation in this study were still implemented in a twin experiment framework, where the model error sources is identified as the estimated parameters. The parameter estimation using the real observational data will be much more complex than that.

One of the major challenges for parameter estimation with real observational data is how to identify of the biased parameters in a CGCM. Aside from the parameter uncertainties, the model bias can be generated in a CGCM due to model structural errors, such as the imperfect dynamical framework and the incomplete understanding for physical processes. It remains a great challenge to identify the sources of the model bias from the candidates of the model structural deficiencies, as well as the large number of model parameters. Hu et al (2010), in their real-data parameter estimation study, pointed out that the parameter estimation using real observations might produce the right answer for the wrong reasons. Furthermore, As discussed in appendix 3.1 in chapter 3, it is not guarantee that a biased parameter can be estimated using EnKF. The parameter estimation is much more challenge for a highly nonlinear parameter

because the EnKF is only a linear filter. We may not be able to extract a signal dominated linear partition from the total model response to the parameter uncertainties.

One major challenge for parameter estimation with real observational data is the deficiency of ensemble spread. The uncertainty generated by the model structural errors cannot be included in a single model ensemble forecast. Therefore, the background uncertainty estimated from the ensemble perturbations usually suffers a negative deficiency when we apply ensemble based filter using real observations. A negatively biased background uncertainty could cause poor filter performance or even for parameter estimation with real observational data and therefore cause parameter estimation failure. One has to inflate the background variance to compromise the uncertainty deficiency and improve filter performance using a state-of-the-art inflation schemes, such as the covariance inflation/relaxation (Zhang et al., 2004), the additive inflation (Hamill and Whitaker, 2005), or the adaptive covariance inflation (Anderson, 2007, 2009). The effect of these inflation schemes on parameter estimation remains to further study.

Some study also use the parameter uncertainty as a carrier for the forecast uncertainty generated by the model structural deficiencies (Hansen, 2002). From this angle, the constant parameters in deterministic model will be regarded as stochastic parameters. Ensemble based parameter estimation techniques can use to estimate the PDFs of stochastic parameters (Hansen and Penland 2007). The implementation of this technique in a CGCM remains to further study.

Reference

- Aksoy, A., F. Zhang and J. W. Nielsen-Gammon, 2006a: Ensemble-based simultaneous state and parameter estimation with MM5. *Geophys. Res. Lett.*, 33, L12801, doi:10.1029/2006GL026186.
- Aksoy, A., F. Zhang and J. W. Nielsen-Gammon, 2006b: Ensemble-based simultaneous state and parameter estimation in a Two-Dimensional Sea-Breeze Model. *Mon. Wea. Rev.*, 134, 2951-2970.
- Anderson, J. L., 1997. The impact of dynamical constraints on the selection of initial conditions for ensemble predictions: low-order perfect model results. *Mon. Wea. Rev.*, 123, 2969-2983
- Anderson, J. L., 2001: An ensemble adjustment Kalman filter for data assimilation. *Mon. Wea. Rev.*, 129, 2884-2902
- Anderson, J. L., 2003: A local least squares framework for ensemble filtering. *Mon. Wea. Rev.*, 131, 634--642.
- Anderson, J. L., 2007: An adaptive covariance inflation error correction algorithm for ensemble filters. *Tellus*, 59A, 210-224.
- Anderson, J. L., 2009: Spatially and temporally varying adaptive covariance inflation for ensemble filters. *Tellus*, 61A, 72-83.
- Anderson, J. L., 2010. A non-Gaussian ensemble filter update for data assimilation. *Mon. Wea. Rev.*, 138,4186–4198.
- Anderson, W. G., A. Gnanadesikan, R. Hallberg, J. Dunne, and B. L. Samuels, 2007: Impact of ocean color on the maintenance of the Pacific cold tongue. *Geophys. Res. Lett.*, 34, L11609, doi: 10.1029/2007GL030100.
- Annan, J. D. and J. C. Hargreaves, 2004: Efficient parameter estimation for a highly chaotic system. *Tellus*, 56A, 520--526.
- Annan, J. D., 2005: Parameter estimation using chaotic time series. *Tellus*, 57A, 709-714
- Annan, J. D., J. C. Hargreaves, N. R. Edwards and R. Marsh, 2005a: Parameter estimation in an intermediate complexity Earth System Model using an ensemble Kalman filter. *Ocean Modeling*, 8, 135--154.
- Annan, J.D., D.J. Lunt, J.C. Hargreaves, and P. J. Valdes, 2005b: Parameter estimation in an atmospheric GCM using the ensemble Kalman filter. *Nonlinear Processes Geophys.* 12, 363-371.
- Asselin, R., 1972. Frequency filter for time integrations. *Mon. Wea. Rev.*, 100, 487-490.
- Ballabrera-Poy, J., R. Murtugudde, R.H. Zhang, and A.J. Busalacchi 2007: Coupled ocean-atmosphere response to seasonal modulation of ocean color: Impact on interannual climate experiments in the tropical Pacific. *J. Clim.*, 20, 353-374.

- Banks, H.T., 1992a: Control and estimation in distributed parameter systems. In: H.T. Banks, Editor, *Frontiers in Applied Mathematics* vol. 11, SIAM, Philadelphia, pp227.
- Banks, H.T., 1992b: Computational issues in parameter estimation and feedback control problems for partial differential equation systems. *Physica D* 60, 226-238.
- Bishop, C. H., B. Etherton, and S. J. Majumdar, 2001. Adaptive sampling with the ensemble transformation kalman filter. part i: theoretical aspects. *Mon. Wea. Rev.*, 129, 420–436
- Burgers, G., P. van Leeuwen, and G. Evensen, 1998. Analysis scheme in the ensemble Kalman filter. *Mon. Wea. Rev.*, 126, 1719–1724.
- Chang, T., S. Zhang, A. Rosati, T. Delworth and W. Stern, 2011: An assessment of oceanic variability for 1960-2010 from the GFDL ensemble coupled data assimilation. *Climate Dynamics*, 40, 775-803.
- Collins, W. and co-authors, 2006: The Community Climate System Model Version 3 (CCSM3). *J. Climate*, 19, 2111-2143
- Davey M. K. et al., 2002: STOIC: a study of coupled model climatology and variability in tropical ocean regions. *Clim. Dyn.*, 18, 403-420
- Derber, J., 1989: A variational continuous assimilation scheme. *Mon. Wea. Rev.*, 117, 2437-2446
- Evensen, G., 1994. Sequential data assimilation with a non-linear quasi-geostrophic model using monte carlo methods to forecast error statistics. *J. Geophys. Res.*, 99(C5), 10 143–10 162.
- Evensen, G., 2003. The ensemble kalman filter: theoretical formulation and practical implementation. *Ocean Dynamics*, 53, 343–367
- Evensen, G., 2007: *Data Assimilation: The Ensemble Kalman Filter*, Springer Press, pp187.
- Fraedrich, K. and C. Ziehmann-Schlumbohm, 1994. Predictability experiments with persistence forecasts in a red-noise atmosphere. *Q. J. R. Meteorol. Soc.*, 120, 387-428
- Frouin, R., D. W. Lingner, C. Gautier, K. S. Baker, and R. C. Smith 1989: A simple analytical formula to compute clear sky total photosynthetically available solar irradiance at the ocean surface. *J. Geophys. Res.*, 94, 9731-9742.
- Gaspari, G., and S. E. Cohn, 1999: Construction of correlation functions in two and three dimensions. *Quart. J. Roy. Meteor. Soc.*, 125, 723–757.
- Hacker, J.P. and C. Snyder, 2005: Ensemble Kalman filter assimilation of fixed screen-height observations in a parameterized PBL. *Mon. Wea. Rev.*, 133, 3260-3275

- Hamill, T. M., and J. S. Whitaker, 2005: Accounting for the error due to unresolved scales in ensemble data assimilation: A comparison of different approaches. *Mon. Wea. Rev.*, 133, 3132–3147.
- Hamill, T. M., J. S. Whitaker, and C. Snyder, 2001. Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter. *Mon. Wea. Rev.*, 129, 2776–2790
- Hamill, T., C. Snyder and R. Morss, 2000: A comparison of probabilistic forecasts from bred, singular-vector and perturbed observation ensembles. *Mon. Wea. Rev.*, 128, 1835-1851
- Hansen, J. A. (2002), Accounting for model error in ensemble-based state estimation and forecasting, *Mon. Weather Rev.*, 130, 2373–2391.
- Hansen, J., and C. Penland (2007), On stochastic parameter estimation using data assimilation, *Physica D*, 230, 88–98.
- Hokanson, Erin P., 2006, The effects of solar penetration on a coupled general circulation model, masters theses, UW-madison, department of atmospheric oceanic sciences.
- Houtekamer, P. L. and H. L. Mitchell, 1998. Data assimilation using an ensemble Kalman filter technique. *Mon. Wea. Rev.*, 126, 796–811.
- Houtekamer, P. L. and H. L. Mitchell, 2001. A sequential ensemble Kalman filter for atmospheric data assimilation. *Mon. Wea. Rev.*, 129, 123–137.
- Houtekamer, P. L., and H. L. Mitchell (1998), Data assimilation using an ensemble Kalman filter technique, *Mon. Weather Rev.*, 126, 796– 811.
- Houtekamer, P. L., G. Pellerin, M. Buehner, M. Charron, L. Spacek, and B. Hansen, 2005: Atmospheric data assimilation with an ensemble Kalman filter: Results with real observations. *Mon. Wea. Rev.*, 133, 604–620.
- Hu, X.-M. , F. Zhang, and J. W. Nielsen-Gammon, 2010: Ensemble-based simultaneous state and parameter estimation for treatment of mesoscale model error: A real-data study, *Geophys. Res. Lett.*, 37, L08802, doi:10.1029/2010GL043017.
- Jacob, R., 1997: Low frequency variability in a simulated atmosphere-ocean system. Ph.D thesis, University of Wisconsin-Madison, 155pp
- Jazwinski, A. H., 1970: *Stochastic Processes and Filtering Theory*. Academic Press, New York, 376pp.
- Kalman, R. and R. Bucy, 1961: New results in linear filtering and prediction theory. *Trans. ASME, Ser. D, J. Basic Eng.*, 83, 95-109.
- Kalman, R. E., 1960. A new approach to linear filtering and prediction problems. *Journal of Basic Engineering*, 82, 34-45.

- Kondrashov, D., C. Sun and M. Ghil, 2008: Data assimilation for a coupled ocean atmosphere model, Part II: Parameter estimation. *Mon. Wea. Rev.*, 136, 5062-5076
- Lawson, G. W. and J. A. Hansen, 2004. Implications of stochastic and deterministic filters as ensemble-based data assimilation methods in varying regimes of error growth. *Mon. Wea. Rev.*, 132, 1966–1981.
- Lei, J. P. Bickel and C. Snyder, 2010. Comparison of Ensemble Kalman Filters under Non-Gaussianity. *Mon. Wea. Rev.*, 138, 1293–1304.
- Lin, J-L., 2007: The double-ITCZ problem in IPCC AR4 coupled GCMs: Ocean–atmosphere feedback analysis. *J. Climate*, 20, 4497-4525.
- Liu, Z., B. Otto-Bliesner, J. Kutzbach, L. Li, C. Shields, 2003: Coupled climate simulations of the evolution of global monsoons in the Holocene. *J. Clim.*, 16, 2472-2490
- Liu, Z., J. Kutzbach and L. Wu, 2000: Modeling climatic shift of El Nino variability in the Holocene. *Geophys. Res. Lett.*, 27, 2265-2268.
- Liu, Z., S. Wu, S. Zhang, Y Liu, and X Rong, September 2013: Ensemble data assimilation in a simple coupled climate model: The role of ocean-atmosphere interaction. *Advances in Atmospheric Sciences*, 30(5), DOI:10.1007/s00376-013-2268-z.
- Liu, Z., Y. Liu, L. Wu and R. Jacob, 2007b: Seasonal and Long-Term Atmospheric Responses to Reemerging North Pacific Ocean Variability: A Combined Dynamical and Statistical Assessment. *J. Climate*, 20, 955–980.
- Liu, Z., Y. Wang, R. Gallimore, F. Gasse, T. Johnson, P. deMenocal, J. Adkins, M. Notaro, I.C. Prentice, J. Kutzbach, R. Jacob, P. Behling, L. Wang and E. Ong, 2007a: Simulating the transient evolution and abrupt change of Northern Africa atmosphere–ocean–terrestrial ecosystem in the Holocene. *Quat. Sci. Rev.*, 26, July 2007, 1818-183, doi:10.1016/j.quascirev.2007.03.002
- Lorenz, E. N., 1963. Deterministic nonperiodic flow. *J. Atmos. Sci.*, 20, 130–141.
- Lorenz, E. N., 1996. Predictability: a problem partly solved. In: *Proceedings’ of the ECMWF seminar on predictability, vol1*, ECMWF, Reading, United Kingdom, 1-18.
- Lorenz, E. N., 2005. Designing chaotic models. *J. Atmos. Sci.*, 62, 1574–1587.
- Mechoso, C. R, et al., 1995: The seasonal cycle over the tropical Pacific in coupled ocean-atmosphere general circulation models. *Mon. Wea. Rev.*, 123, 2825-2838
- Mitchell, H. L. and P. L. Houtekamer and G. Pellerin, 2002. Ensemble size, balance, and model-error representation in an ensemble Kalman filter. *Mon. Wea. Rev.*, 130, 2791–2808.
- Mitchell, H. L. and P. L. Houtekamer, 2009. Ensemble Kalman Filter Configurations and Their Performance with the Logistic Map. *Mon. Wea. Rev.*, 137, 4325-4343.

- Murtugudde, R., J. Beauchamp, and A. Busalaccki 2002: Effects of penetrative radiation on the upper tropical ocean circulation. *J. Clim.*, 15, 470-486.
- Nakamoto, S., S. Prasanna Kumar, J. M. Oberhuber, J. Ikshizaka, K. Muneyama, and R. Frouin 2001: Response of the equatorial Pacific to chlorophyll pigment in a mixed layer isopycnal ocean general circulation model. *Geophys. Res. Lett.*, 28, 2021-2024.
- Navon, I. M., 1997: Practical and theoretical aspects of adjoint parameter estimation and identifiability in meteorology and oceanography. *Dyn. Atmos. Ocn.*, 27, 55-79
- Nielsen-Gammon, J. W., X.-M. Hu, F. Zhang, and J. E. Pleim, 2010: Evaluation of Planetary Boundary Layer Scheme Sensitivities for the Purpose of Parameter Estimation, *Mon. Wea. Rev.*, 138, 3400-3417.
- Ohlmann, J. C., D. A. Siegel, and C. D. Mobley 2000: Ocean radiant heating. Part I; Optical influences. *J. Phys. Oceanogr.*, 30, 1833-1848.
- Pacanowski, R. C., and S. G.H. Philander, 1981: Parameterization of vertical mixing in numerical models of tropical oceans. *J. Phys. Oceanogr.*, 11, 1443-1451.
- Peters, H., M. C. Gregg and J. M. Tool 1988: On the parameterization of equatorial turbulence. *J. Geophys. Res.*, 92, 1199-1218.
- Pham, D. T., 2001. Stochastic methods for sequential data assimilation in strongly nonlinear systems. *Mon. Wea. Rev.*, 129, 1194–1207.
- Ridgwell, A., J. C. Hargreaves, N.R. Edwards, J.D. Annan, T. M. Lenton, R. Marsh, A. Yool and A. Watson, 2007: Marine geochemical data assimilation in an efficient Earth System Model of global biogeochemical cycling. *Biogeosciences*, 4, 87-104
- Ridgwell, A., J. C. Hargreaves, N.R. Edwards, J.D. Annan, T. M. Lenton, R. Marsh, A. Yool and A. Watson, 2007: Marine geochemical data assimilation in an efficient Earth System Model of global biogeochemical cycling. *Biogeosciences*, 4, 87-104
- Robert, A., 1969. The integration for a spectrum model of the atmosphere by the implicit method. *Proc. WOM/IUGG symposium on NWP*, Japan meteorological society, Tokyo, Japan, 19-24.
- Sacher, W. and P. Bartello, 2008. Sampling errors in ensemble Kalman filtering. Part i: theory. *Mon. Wea. Rev.*, 136, 3035–3049.
- Sakov, P. and P. R. Oke, 2008. Implications of the form of the ensemble transformation in the ensemble square root filters. *Mon. Wea. Rev.*, 136, 1042–1053.
- Schneider, E., and Z. Zhu 1998: Sensitivity of the simulated annual cycle of sea surface temperature in the equatorial Pacific to sunlight penetration. *J. Clim.*, 11, 1932-1950.

- Smith, R. C. and K. S. Baker 1978: The bio-optical state of ocean waters and remote sensing. *Limnol. Oceanogr.*, 23, 247-259.
- Tippett, M. K., J. L. Anderson, C. H. Bishop, T. M. Hamill, and J. S. Whitaker, 2003. Ensemble square root filters. *Mon. Wea. Rev.*, 131, 1485–1490.
- Tong, M. and M. Xue, 2008a: Simultaneous estimation of microphysical parameters and atmospheric state with simulated Radar data and ensemble square root Kalman filter. Part I: Sensitivity analysis and parameter identifiability. *Mon. Wea. Rev.*, 136, 1630-1648.
- Tong, M. and M. Xue, 2008b: Simultaneous estimation of microphysical parameters and atmospheric state with simulated Radar data and ensemble square root Kalman filter. Part II: Parameter estimation experiments. *Mon. Wea. Rev.*, 136, 1649-1668.
- Whitaker, J. S. and T. M. Hamill, 2002. Ensemble data assimilation without perturbed observations. *Mon. Wea. Rev.*, 130, 1913–1924.
- Wu, X., S. Zhang, Z. Liu, A. Rosati, and T. Delworth, 2012b: A study of impact of the geographic dependence of observing system on parameter estimation with an intermediate coupled model. *Climate dynamics*. Doi:10.1007/s00382-012-1385-1.
- Wu, X., S. Zhang, Z. Liu, A. Rosati, T. Delworth and Y. Liu, 2012a: Impact of Geographic Dependent Parameter Optimization on Climate Estimation and Prediction: Simulation with an Intermediate Coupled Model. *Mon. Wea. Rev.*, 140, 3956–3971.
- Wu, L, Z. Liu, R. Gallimore, R. Jacob, D. Lee, and Y. Zhong, 2003: A coupled modeling study of Pacific decadal variability: The Tropical Mode and The North Pacific Mode. *J. Climate*, 16, 1101-1120
- Yang, X., A. Rosati, S. Zhang, T. Delworth, R. Gudgel, R. Zhang, G. Vecchi, W. Anderson, Y. Chang, T. DelSole, K. Dixon, R. Msadek, W. Stern, A. Wittenberg and F. Zeng, 2013: A predictable AMO-like pattern in GFDL's fully-coupled ensemble initialization and decadal forecasting system. *J. Climate*, 26, 650-661.
- Zhang, F., C. Snyder, and J. Sun, 2004: Impact of initial estimate and observation availability on convective-scale data assimilation with an ensemble Kalman filter. *Mon. Wea. Rev.*, 132, 1238–1253.
- Zhang, S. 2011b: A study of impacts of coupled model initial shocks and state-parameter optimization on climate predictions using a simple pycnocline prediction model. *Journal of Climate*, 24(23), 6210–6226. doi: <http://dx.doi.org/10.1175/JCLI-D-10-05003.1>
- Zhang, S. and J. L. Anderson, 2003: Impact of spatially and temporally varying estimates of error covariance on assimilation in a simple atmospheric model. *Tellus*, 55A, 126-147.
- Zhang, S., J. Harrison, A. Rosati and A. Wittenberg, 2007: System design and evaluation of coupled ensemble data assimilation for global oceanic climate studies. *Mon. Wea. Rev.*, 135, 3541-3564

- Zhang, S., 2011, Impact of observation optimized model parameters on decadal predictions: Simulation with a simple pycnocline prediction model, *Geophys. Res. Lett.*, 38, L02702.
- Zhang, S., 2011a: Impact of Observation-Optimized Model Parameters on Decadal Predictions: Simulation with a Simple Pycnocline Prediction Model. *Geophys. Res. Lett.*, 38, L02702, doi:10.1029/2010GL046133.
- Zhang, S., A. Rosati, M. J. Harrison, R. Gudgel and W. Stern, 2008: GFDL's coupled ensemble data assimilation system, 1980-2006 coupled reanalysis and its impact on ENSO forecast. Preprint wcrp3, Jan. 27 - Feb. 2, 2008, Tokyo http://wcrp.ipsl.jussieu.fr/Workshops/Reanalysis2008/Documents/G4-434_ea.pdf
- Zhang, S., J. L. Anderson, A. Rosati, M. Harrison, S. P. Khare, A. Wittenburg, 2004. Multiple time level adjustment for data assimilation. *Tellus*, 56A, 2-15.
- Zhang, S., M. J. Harrison, A. Rosati and A. T. Wittenberg, 2007: System design and evaluation of coupled ensemble data assimilation for global oceanic climate studies. *Mon. Wea. Rev.*, 135, 3541-3564.
- Zhang, S., Y. Chang, X. Yang and A. Rosati, 2014: Balanced and Coherent Climate Estimation by Combining Data with a Biased Coupled Model. *J. Climate*, 27(3) , 1302-1314.
- Zhang, S., Z. Liu, A. Rosati and T. Delworth, 2012: A Study of Enhancive Parameter Correction with Coupled Data Assimilation for Climate Estimation and Prediction Using a Simple Coupled Model. *Tellus A*, 64, 10963, DOI:10.3402/tellusa.v64i0.10963.
- Zhu, Y., and I. M. Navon, 1999: Impact of parameter estimation on the performance of the FSU global spectral model using its full-physics adjoint. *Mon. Wea. Rev.*, 127, 1497-1517
- Zou, X., I.M. Navon and F.X. LeDimet, 1992: An optimal nudging data assimilation scheme using parameter estimation. *Quart. J. Roy. Meteor. Soc.*, 118, 1163-1186